

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 5, 2004, 13:25:43 ; Search time 1.09687 Seconds  
(without alignments)  
3164.197 Million cell updates/sec

Title: US-09-743-684A-19  
Perfect score: 63  
Sequence: 1 GDVQCDCIQMW 11

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 1017041 seqs, 315518202 residues

Total number of hits satisfying chosen parameters: 1017041

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : SPTREMBL\_25.\*  
1: sp\_archaea.\*  
2: sp\_bacteria.\*  
3: sp\_fungi.\*  
4: sp\_human.\*  
5: sp\_invertebrate.\*  
6: sp\_mammal.\*  
7: sp\_mhc.\*  
8: sp\_organelle.\*  
9: sp\_phase.\*  
10: sp\_plant.\*  
11: sp\_rodent.\*  
12: sp\_virus.\*  
13: sp\_vertebrate.\*  
14: sp\_unclassified.\*  
15: sp\_rvirus.\*  
16: sp\_bacteriap.\*  
17: sp\_archaeap.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	50	79.4	520	13 Q8UVZ4	Q8uvz4 brachydanio
2	50	79.4	522	13 Q9DG82	Q9dg82 brachydanio
3	47	74.6	554	11 Q8BFQ1	Q8bfq1 mus musculus
4	45	71.4	402	11 Q8CLN0	Q8cln0 mus musculus
5	45	71.4	449	11 Q8BJV5	Q8bjv5 mus musculus
6	45	71.4	525	11 Q8C1C1	Q8clc1 mus musculus
7	44	69.8	221	15 Q8DV19	Q8dvi9 streptococc
8	42	66.7	512	13 Q7S770	Q7sy70 xenopus lae
9	42	66.7	957	10 Q94GY5	Q94gy5 oryza sativ
10	41	65.1	465	11 Q9P9K8	Q9p9k8 methanosarc
11	41	65.1	480	17 Q8TRG4	Q8trg4 methanosarc
12	40	63.5	46	16 O07606	O07606 bacillus su
13	40	63.5	155	6 O77634	O77634 sus scrofa
14	40	63.5	181	16 Q7VF78	Q7vft8 helicobacte
15	40	63.5	184	2 Q9ANV3	Q9anv3 lactobacill
16	40	63.5	328	10 Q8GYB6	Q8gyb6 arabidopsis

17	40	63.5	545	12 Q9WAT6	Q9wat6 chuzan viru
18	40	63.5	552	12 Q41929	Q41929 murid herpe
19	40	63.5	972	5 Q9XZ08	Q9xz08 drosophila
20	40	63.5	1356	10 Q9SH77	Q9sh77 arabidopsis
21	39	61.9	124	12 Q8JKN4	Q8jkn4 heliothis z
22	39	61.9	167	12 Q8VBP4	Q8vbp4 thogoto vir
23	39	61.9	167	12 Q8VBP5	Q8vbp5 thogoto vir
24	39	61.9	257	3 Q871L9	Q871l9 neurospora
25	39	61.9	320	12 Q86517	Q86517 rotavirus
26	39	61.9	388	5 Q9N652	Q9n652 drosophila
27	39	61.9	388	5 Q9NGG6	Q9ngg6 drosophila
28	39	61.9	388	5 Q9N656	Q9n656 drosophila
29	39	61.9	388	5 Q9N6D3	Q9n6d3 drosophila
30	39	61.9	666	10 Q9LDM5	Q9ldm5 arabidopsis
31	39	61.9	753	16 Q8E6M8	Q8e6m8 streptococc
32	39	61.9	753	16 Q8E174	Q8e174 streptococc
33	39	61.9	790	5 Q9GSF3	Q9gsf3 podocoryne
34	39	61.9	808	4 Q96JN3	Q96jn3 homo sapien
35	39	61.9	1235	5 Q9VYL6	Q9vyl6 drosophila
36	39	61.9	1235	5 Q9GV19	Q9gv19 drosophila
37	39	61.9	1257	5 Q9GV18	Q9gv18 drosophila
38	39	61.9	1257	5 Q8IR79	Q8ir79 drosophila
39	38	60.3	117	5 Q9V8F5	Q9v8f5 drosophila
40	38	60.3	150	16 Q81X24	Q81x24 bacillus an
41	38	60.3	216	13 Q7SKM9	Q7skm9 brachydanio
42	38	60.3	221	4 Q9NPF0	Q9npf0 homo sapien
43	38	60.3	221	11 Q8C8M1	Q8c8m1 mus musculu
44	38	60.3	228	16 Q8RCE2	Q8rce2 thermonaer
45	38	60.3	277	11 P70361	P70361 mus musculu

## ALIGNMENTS

## RESULT 1

Q8UVZ4 ID Q8UVZ4 PRELIMINARY; PRT; 520 AA.  
AC Q8UVZ4;  
DT 01-MAR-2002 (TrEMBLrel. 20, Created)  
DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)  
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
DE Lysosomal cofactor/neurotrophic factor prosaposin.  
OS Brachydanio rerio (Zebrafish) (Danio rerio).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;  
OC Cyprinidae; Danio.  
OX NCBI\_TaxID=7955;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Seo H.-C., Lie O., Fjose A., O'Brien J.S., Kishimoto Y.;  
RT "Cloning, expression and promoter analysis of zebrafish prosaposin.";  
RL Submitted (JUN-2000) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AF276996; AAL54381.1;  
DR GO; GO:0005764; C:lysosome; IEA.  
DR GO; GO:0006665; P:sphingolipid metabolism; IEA.  
DR InterPro; IPR003119; SapA.  
DR InterPro; IPR007856; SapB\_1.  
DR InterPro; IPR008138; SapB\_2.  
DR InterPro; IPR008140; SapB\_sub.  
DR InterPro; IPR008373; Saposin.  
DR InterPro; IPR008139; SaposinB.  
DR Pfam; PF02199; SAPA; 2.  
DR Pfam; PF05184; SapB\_1; 3.  
DR Pfam; PF03489; SapB\_2; 4.  
DR PRINTS; PR01797; SAPOSIN.  
DR PRODOM; PD001732; SapB\_sub; 3.  
DR SMART; SM00162; SAPA; 2.  
DR SMART; SM00118; SAPB; 4.  
SQ SEQUENCE 520 AA; 57431 MW; F9E620F84BA41CB5 CRC64;

Query Match 79.4%; Score 50; DB 13; Length 520;  
Best Local Similarity 63.6%; Pred. No. 0.62;  
Matches 7; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

```
QY 1 GDVQCDCIQMV 11
Db 186 GDVQCDCVTFI 196

RESULT 2
Q9DG82
ID Q9DG82 PRELIMINARY; PRT; 522 AA.
AC Q9DG82;
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Prosaposin.
GN PSAP.
OS Brachydanio rerio (Zebrafish) (Danio rerio).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
OC Cyprinidae; Danio.
OX NCBI_TaxID=7955;
RN [1]
RP SEQUENCE FROM N.A.
RA Altman N., Horowitz M.;
RT "The zebrafish prosaposin cDNA.";
RL Submitted (NOV-1998) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF108655; AAG32919.1; -.
DR ZFIN; ZDB-GENE-020108-1; psap.
DR GO; GO:0005764; C:lysosome; IEA.
DR GO; GO:0006665; P:sphingolipid metabolism; IEA.
DR InterPro; IPR003119; Sapa.
DR InterPro; IPR007856; SappB_1.
DR InterPro; IPR008138; SappB_2.
DR InterPro; IPR008140; SappB_sub.
DR InterPro; IPR008373; Saposin.
DR InterPro; IPR008139; SaposinB.
DR Pfam; PF02199; SAPA; 2.
DR Pfam; PF05184; SappB_1; 4.
DR Pfam; PF03489; SappB_2; 3.
DR PRINTS; PR01797; SAPOSIN.
DR ProDom; PD001732; SappB_sub; 3.
DR SMART; SM00162; SAPA; 2.
DR SMART; SM00118; SAPPB; 4.
SQ SEQUENCE 522 AA; 57671 MW; D3C15A305725C1CD CRC64;

Query Match 79.4%; Score 50; DB 13; Length 522;
Best Local Similarity 63.6%; Pred. No. 0.62;
Matches 7; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 GDVQCDCIQMV 11
Db 186 GDVQCDCVTFI 196

RESULT 3
Q8BFQ1
ID Q8BFQ1 PRELIMINARY; PRT; 554 AA.
AC Q8BFQ1;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Prosaposin.
GN PSAP.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RA STRAIN=NOD; TISSUE=Kidney, and Thymus;
RL MEDLINE=22354683; PubMed=12466851;
DR EMBL; AK009408; BAC25258.1; -.
DR MGD; MGI:1924193; 2310020A21Rik.
DR GO; GO:0005764; C:lysosome; IEA.
DR GO; GO:0006665; P:sphingolipid metabolism; IEA.
DR InterPro; IPR003119; Sapa.
DR InterPro; IPR007856; SappB_1.
DR InterPro; IPR008138; SappB_2.
DR InterPro; IPR008140; SappB_sub.
DR InterPro; IPR008373; Saposin.
DR InterPro; IPR008139; SaposinB.
DR Pfam; PF02199; SAPA; 1.
DR Pfam; PF05184; SappB_1; 3.
DR Pfam; PF03489; SappB_2; 3.
DR PRINTS; PR01797; SAPOSIN.
DR ProDom; PD001732; SappB_sub; 2.
DR SMART; SM00162; SAPA; 1.
DR SMART; SM00118; SAPPB; 3.
KW Hypothetical protein.
RP "Analysis of the mouse transcriptome based on functional annotation of
```

```

Query Match      71.4%; Score 45; DB 11; Length 402;
Best Local Similarity 54.5%; Pred. NO. 4;
Matches 6; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 1 GDVQCDCIQMV 11
DB 64 GAVCHDCVQLI 74

RESULT 5
Q8BJV5 PRELIMINARY; PRT; 449 AA.
AC Q8BUV5;
DT 01-MAR-2003 (TReMBLrel. 23, Created)
DT 01-MAR-2003 (TReMBLrel. 23, Last sequence update)
DT 01-OCT-2003 (TReMBLrel. 25, Last annotation update)
DE Hypothetical saposin A-type domain/saposin type B containing protein.
GN 2310020A21RIK.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OC NCBI_TaxID=10090;
RN [1]_TaxID=10090;
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Skin;
RX MEDLINE=22354683; PubMed=12466851;
RA The FANTOM Consortium;
RT "Analysis of the mouse transcriptome based on functional annotation of
RT 60,770 full-length cDNAs."
RL Nature 420:563-573(2002).
DR EMBL; AK028455; BAC25961.1; -.
DR MGD; MGI:1924193; 2310020A21RIK.
DR GO; GO:0005764; C:lysosome; IEA.
DR GO; GO:0006665; P:sphingolipid metabolism; IEA.
DR InterPro; IPR003119; Sapa.
DR InterPro; IPR007856; SapB_1.
DR InterPro; IPR008138; SapB_2.
DR InterPro; IPR008140; SapB_sub.
DR InterPro; IPR008373; Saposin.
DR Pfam; PF02199; SAPA; 2.
DR Pfam; PF05184; SapB_1; 4.
DR Pfam; PF03489; SapB_2; 4.
DR PRINTS; PR01797; SAPOSIN.
DR ProDom; PD001732; SapB_sub; 3.
DR SMART; SM00162; SAPA; 2.
DR SMART; SM00118; SAPB; 4.
KW Hypothetical protein.
SQ SEQUENCE 449 AA; 48705 MW; 377CC42A475B292F CRC64;

Query Match      71.4%; Score 45; DB 11; Length 449;
Best Local Similarity 54.5%; Pred. NO. 4.4;
Matches 6; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 1 GDVQCDCIQMV 11
DB 187 GAVCHDCVQLI 197

RESULT 6
Q8C1C1 PRELIMINARY; PRT; 525 AA.
AC Q8C1C1;
DT 01-MAR-2003 (TReMBLrel. 23, Created)
DT 01-MAR-2003 (TReMBLrel. 23, Last sequence update)
DT 01-OCT-2003 (TReMBLrel. 25, Last annotation update)
DE Hypothetical saposin A-type domain/saposin type B containing protein.
GN 2310020A21RIK.
OS Mus musculus (Mouse).

```

```

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OC NCBI_TaxID=10090;
RN [1]_TaxID=10090;
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Skin;
RX MEDLINE=22354683; PubMed=12466851;
RA The FANTOM Consortium;
RT "Analysis of the mouse transcriptome based on functional annotation of
RT 60,770 full-length cDNAs."
RL Nature 420:563-573(2002).
DR EMBL; AK028455; BAC25961.1; -.
DR MGD; MGI:1924193; 2310020A21RIK.
DR GO; GO:0005764; C:lysosome; IEA.
DR GO; GO:0006665; P:sphingolipid metabolism; IEA.
DR InterPro; IPR003119; Sapa.
DR InterPro; IPR007856; SapB_1.
DR InterPro; IPR008138; SapB_2.
DR InterPro; IPR008140; SapB_sub.
DR InterPro; IPR008373; Saposin.
DR Pfam; PF02199; SAPA; 2.
DR Pfam; PF05184; SapB_1; 4.
DR Pfam; PF03489; SapB_2; 4.
DR PRINTS; PR01797; SAPOSIN.
DR ProDom; PD001732; SapB_sub; 3.
DR SMART; SM00162; SAPA; 2.
DR SMART; SM00118; SAPB; 4.
KW Hypothetical protein.
SQ SEQUENCE 525 AA; 57350 MW; 374F6050CDC4D223 CRC64;

Query Match      71.4%; Score 45; DB 11; Length 525;
Best Local Similarity 54.5%; Pred. NO. 5;
Matches 6; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 1 GDVQCDCIQMV 11
DB 187 GAVCHDCVQLI 197

RESULT 7
Q8DV19 PRELIMINARY; PRT; 221 AA.
AC Q8DV19;
DT 01-MAR-2003 (TReMBLrel. 23, Created)
DT 01-MAR-2003 (TReMBLrel. 23, Last sequence update)
DT 01-OCT-2003 (TReMBLrel. 25, Last annotation update)
DE Putative late competence protein.
GN SMU.499.
OS Streptococcus mutans.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OC NCBI_TaxID=1309;
RN [1]_TaxID=1309;
RP SEQUENCE FROM N.A.
RC STRAIN=UA159 / ATCC 700610 / Serotype C;
RX MEDLINE=22295063; PubMed=12397186;
RA Ajdic D., McShan W.M., McLaughlin R.E., Savic G., Chang J.,
RA Carson M.B., Primeaux C., Tian R., Kenton S., Jia H., Lin S., Qian Y.,
RA Li S., Zhu H., Najjar F., Lai H., White J., Roe B.A., Ferretti J.J.;
RT "Genome sequence of Streptococcus mutans UA159, a cariogenic dental
RT pathogen."
RL Proc. Natl. Acad. Sci. U.S.A. 99:14434-14439(2002).
DR EMBL; AE014895; AAN58244.1; -.
DR GO; GO:0005489; F:electron transporter activity; IEA.
DR GO; GO:0006118; P:electron transport; IEA.
DR InterPro; IPR000345; CytC heme BS.
DR PROSITE; PS00190; CYTOCHROME_C; 2.
KW Complete proteome.
SQ SEQUENCE 221 AA; 26017 MW; C3C68D1C6A7F697D CRC64;

Query Match      69.8%; Score 44; DB 16; Length 221;

```

Best Local Similarity 85.7%; Pred. No. 3.5; Mismatches 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 GDVQCQC 7  
DB 54 GDICQC 60

## RESULT 8

Q7SY70 Q7SY70 PRELIMINARY; PRT; 512 AA.  
AC Q7SY70;  
DT 01-OCT-2003 (TrEMBLrel. 25, Created)  
DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)  
DE Hypothetical protein.  
OS Xenopus laevis (African clawed frog).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae; Pipidae;  
OC Xenopodinae; Xenopus.  
OX NCBI\_TaxID=8355;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Whole;  
RX MEDLINE=22341132; PubMed=12454917;  
RA Klein S.L., Strausberg R.L., Wagner L., Pontius J., Clifton S.W.,  
RA Richardson P.;  
RT "Genetic and genomic tools for Xenopus research: The NIH Xenopus  
RT initiative";  
RL Dev. Dyn. 225:384-391(2002).  
RN [2]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Whole;  
RX MEDLINE=22388257; PubMed=12477932;  
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,  
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,  
RA Altshul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,  
RA Hopkins R.F., Jordan H., Moore T.I., Max S.I., Wang J., Hsieh F.,  
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,  
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,  
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,  
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullany S.J.,  
RA Boeak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,  
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,  
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,  
RA Fahey J., Helton E., Kettman A., Madan A., Rodriguez S., Sanchez A.,  
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,  
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,  
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,  
RA Krzywinski M.I., Skalska U., Smailus D.E., Schnerch A., Schein J.E.,  
RA Jones S.J., Marra M.A.;  
RT "Generation and initial analysis of more than 15,000 full-length human  
RT and mouse cDNA sequences";  
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).  
RN [3]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Whole;  
RA Klein S., Strausberg R.;  
RL Submitted (JUL-2003) to the EMBL/GenBank/DBJ databases.  
DR EMBL; BC054988; AAK54988.1; -;  
KW Hypothetical protein.  
SQ SEQUENCE 512 AA; 57555 MW; 57CFA5E2093F6FB1 CRC64;

Query Match 66.7%; Score 42; DB 13; Length 512;  
Best Local Similarity 45.5%; Pred. No. 17;  
Matches 5; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 1 GDVQCQCQV 11  
DB 192 GDICNCTTLLI 202

## RESULT 9

Q94GY5 Q94GY5 PRELIMINARY; PRT; 957 AA.

AC Q94GY5;  
DT 01-DEC-2001 (TrEMBLrel. 19, Created)  
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)  
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
DE Hypothetical protein.  
OS Oryza sativa (Rice).  
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;  
OC Ehrhartoidae; Oryzaceae; Oryza.  
OX NCBI\_TaxID=4530;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=cv. Nipponbare;  
RA Buell C.R., Yuan Q., Ouyang S., Moffat K.S., Hill J.N., Gansberger K.,  
RA Brenner M., Burgess S., Hance M., Shvartsbeyn M., Teitrit T.,  
RA Riggs F., Heiao J., Zismann V., Blunt S., Pal G., VanAken S.E.,  
RA Unterback T.R., Feldblyum T.V., Quackenbush J., Salzberg S.L.,  
RA White O., Fraser C.M.;  
RT "Oryza sativa chromosome 3 BAC OSJNBA0018H01 genomic sequence.";  
RL Submitted (AUG-2001) to the EMBL/GenBank/DBJ databases.  
RX EMBL; AC087181; AAK38494.1; -;  
DR Gramene; Q94GY5; -;  
DR InterPro; IPR007656; DUF593.  
DR Pfam; PF04576; DUF593; 1.  
KW Hypothetical protein.  
SQ SEQUENCE 957 AA; 105813 MW; 6C569D73B33C3C38 CRC64;

Query Match 66.7%; Score 42; DB 10; Length 957;  
Best Local Similarity 62.5%; Pred. No. 30;  
Matches 5; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 1 GDVQCQC 8  
DB 113 GDICQDCV 120

## RESULT 10

Q9P9K8 Q9P9K8 PRELIMINARY; PRT; 465 AA.  
AC Q9P9K8;  
DT 01-OCT-2000 (TrEMBLrel. 15, Created)  
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)  
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)  
DE Putative vanadium dinitrogenase alpha subunit.  
GN VNF.  
OS Methanosaerina barkeri.  
OC Archaea; Euryarchaeota; Euryarchaeota orders incertae sedis;  
OC Methanosaerinales; Methanosaerinae; Methanosaerina.  
OX NCBI\_TaxID=2208;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=ATCC 43241 / DSM 1538 / 227;  
RX MEDLINE=20270156; PubMed=10809706;  
RA Chien Y.T., Auerbuch V., Brabban A.D., Zinder S.H.;  
RT "Analysis of genes encoding an alternative nitrogenase in the archaeon  
RT Methanosaerina barkeri 227.";  
RL J. Bacteriol. 182:3247-3253(2000).  
DR EMBL; AF254784; AAF72180.1; -;  
DR HSSP; P00467; 1MIO.  
DR GO; GO:0016734; F:molybdenum-iron nitrogenase activity; IEA.  
DR GO; GO:0016491; F:oxidoreductase activity; IEA.  
DR GO; GO:0006118; P:electron transport; IEA.  
DR GO; GO:0009339; P:nitrogen fixation; IEA.  
DR InterPro; IPR000318; Nitrogenase comp.  
DR InterPro; IPR005974; Oxidized nitro; 1.  
DR InterPro; IPR000510; Oxidized nitro; 1.  
DR TIGRFAms; TIGR01284; a1c nitrog alph; 1.  
DR PROSITE; PS00699; NITROGENASE\_1; 1.  
SQ SEQUENCE 465 AA; 52841 MW; B9312F97C876D9A3 CRC64;



Query Match 65.1%; Score 41; DB 1; Length 465;  
 Best Local Similarity 63.6%; Pred. No. 24;  
 Matches 7; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Qy 1 GDVQCDCIQMV 11  
 | | | | |  
 Db 59 GGVINDCIQMI 69

## RESULT 11

Q8TRG4 PRELIMINARY; PRT; 480 AA.  
 AC Q8TRG4;  
 DT 01-JUN-2002 (TrEMBLrel. 21, Created)  
 DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)  
 DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)  
 DE Nitrogenase, subunit alpha.  
 GN VNF0 OR M1216.  
 OS Methanosaerina acetivorans.  
 OC Archaea; Euryarchaeota; Euryarchaeota orders incertae sedis;  
 OC Methanosarcinales; Methanosarcinaceae; Methanosarcina.  
 OX NCBI\_TaxID=2214;  
 EN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=C2A / ATCC 35395 / DSM 2834;  
 RX MEDLINE=2192760; PubMed=11932238;  
 RA Galagan J.E., Nusbaum C., Roy A., Endrizzi M.G., MacDonald P.,  
 RA FitzHugh W., Calvo S., Engels-Thmann N., Dearellano K., Johnson R.,  
 RA Allen N., Naylor J., Stange-Romann K., Talarmin J., Tirrell A., Ye W.,  
 RA Linton L., McEwan P., McKernan K., Talarmin J., Tirrell A., Ye W.,  
 RA Zimmer A., Barber R.D., Cann I., Graham D.E., Graham D.A., Guss A.M.,  
 RA Hederich R., Ingram-Smith C., Kuetner H.C., Krzycki J.A.,  
 RA Leigh J.A., Li W., Liu J., Mukhopadhyay B., Reeve J.N., Smith K.,  
 RA Springer T.A., Unayam L.A., White O., White R.H., de Macario E.C.,  
 RA Ferry J.G., Jarrell K.F., Jing H., Macario A.J.L., Paulsen I.,  
 RA Pritchett M., Sowers K.R., Swanson R.V., Zinder S.H., Lander E.,  
 RA Metcalf W.W., Birren B.;  
 RA "The genome of Methanosaerina acetivorans reveals extensive metabolic  
 RT and physiological diversity";  
 RL Genome Res. 12:532-542(2002).  
 DR EMBL; AE010789; AAM04635.1; --  
 DR HSSP; P07328; 1MIN.  
 DR GO; GO:0016734; F:molybdenum-iron nitrogenase activity; IEA.  
 DR GO; GO:0016491; F:oxidoreductase activity; IEA.  
 DR GO; GO:0006118; P:electron transport; IEA.  
 DR GO; GO:0009399; P:nitrogen fixation; IEA.  
 DR InterPro; IPR000318; Nitrogenase compl.  
 DR InterPro; IPR005974; Nitrog\_n\_alpha.  
 DR InterPro; IPR000510; Oxidred\_nitrognasel.  
 DR Pfam; PF00148; Oxidored\_nitro; 1.  
 DR TIGRFAMs; TIGR01284; alt\_nitroq\_alph; 1.  
 DR PROSITE; PS00699; NITROGENASE\_1\_1; 1.  
 KW Complete proteome.  
 SQ SEQUENCE 480 AA; 54598 MW; E8D57106E82474D CRC64;

Query Match 65.1%; Score 41; DB 17; Length 480;  
 Best Local Similarity 63.6%; Pred. No. 24;  
 Matches 7; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Qy 1 GDVQCDCIQMV 11  
 | | | | |  
 Db 59 GGVINDCIQMI 69

## RESULT 12

O07606 PRELIMINARY; PRT; 46 AA.  
 ID O07606;  
 AC O07606;  
 DT 01-JUL-1997 (TrEMBLrel. 04, Created)  
 DT 01-JAN-1998 (TrEMBLrel. 05, Last sequence update)  
 DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)  
 DE Hypothetical protein ynfH.  
 GN YNFH.

OS Bacillus subtilis.  
 OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.  
 OX NCBI\_TaxID=1423;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=168;  
 RA Noback M.A., Terpstra P., Holsappel S., Venema G., Bron S.;  
 RL Submitted (JUN-1997) to the EMBL/GenBank/DBJ databases.  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=168;  
 RX MEDLINE=98044033; PubMed=9384377;  
 RA Kunst F., Ogasawara N., Moszer I., Albertini A.M., Alloni G.,  
 RA Azevedo V., Bertero M.G., Bessieres P., Bolotin A., Borchert S.,  
 RA Borries R., Boursier L., Brans A., Braun M., Brignell S.C., Bron S.,  
 RA Brouillet S., Bruschi C.V., Caldwell B., Capuano V., Carter N.M.,  
 RA Choi S.K., Codani J.J., Conterstorf I.F., Cummings N.J., Daniel R.A.,  
 RA Denizot F., Devine K.M., Dusterhoft A., Ehrlich S.D., Emerson P.T.,  
 RA Entian K.D., Errington J., Fabret C., Ferrari E., Foulger D.,  
 RA Fritz C., Fujita M., Fujita Y., Fuma S., Galizzi A., Galleron N.,  
 RA Ghim S.Y., Glaser P., Goffeau A., Golightly E.J., Grandi G.,  
 RA Guisepi G., Guy B.J., Haga K., Haelegh J., Harwood C.R., Henaut A.,  
 RA Hilbert H., Holsappel S., Hosono S., Hulio M.F., Itaya M., Jones L.,  
 RA Joris B., Karamata D., Kaahara Y., Klaerr-Blanchard M., Klein C.,  
 RA Kobayashi Y., Koetter P., Koningsstein G., Krogh S., Kumano M.,  
 RA Kurita K., Lapidus A., Lardinois S., Lauber J., Lazarevic V.,  
 RA Lee S.M., Levine A., Liu H., Masuda S., Mauel C., Medigue C.,  
 RA Medina N., Mellado R.P., Mizuno M., Moestl D., Nakai S., Noback M.,  
 RA Noone D., O'Reilly M., Ogawa K., Ogiwara A., Oudega B., Park S.H.,  
 RA Parro V., Pohl T.M., Portetelle D., Porwollik S., Prescott A.M.,  
 RA Presecan E., Pujic P., Purnelle B., Rapoport G., Rey M., Reynolds S.,  
 RA Rieger M., Rivolta C., Rocha E., Roche B., Rose M., Sadaie Y.,  
 RA Sato T., Scanlan E., Schleich S., Schroeter R., Scoffone F.,  
 RA Sekiguchi J., Sekowska A., Seror S.J., Serror P., Shin B.S., Soldo B.,  
 RA Srokin A., Tacconi E., Takagi T., Takahashi H., Takemaru K.,  
 RA Takeuchi M., Tamakoshi A., Tanaka T., Terpetra P., Tognoni A.,  
 RA Tosato V., Uchiyama S., Vandenbol M., Vannier F., Vassarotti A.,  
 RA Viari A., Wambutt R., Wedler E., Wedler H., Weitzenecker T.,  
 RA Winters P., Wipat A., Yamamoto H., Yamane K., Yasumoto K., Yata K.,  
 RA Yoshida K., Yoshikawa H.F., Zumatein E., Yoshikawa H., Danchin A.;  
 RT "The complete genome sequence of the Gram-positive bacterium Bacillus  
 RT subtilis";  
 RL Nature 390:249-256(1997).  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=168;  
 RA Kunst F., Ogasawara N., Yoshikawa H., Danchin A.;  
 RL Submitted (NOV-1997) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; Y14083; CAA74529.1; --  
 DR EMBL; Z99109; CAB12863.1; --  
 DR PIR; E69830; E69830.  
 KW Hypothetical protein; Complete proteome.  
 SQ SEQUENCE 46 AA; 5294 MW; AF80B4AA71C18061 CRC64;

Query Match 63.5%; Score 40; DB 16; Length 46;  
 Best Local Similarity 50.0%; Pred. No. 4.6;  
 Matches 5; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GDVQCDCIQM 10  
 | | | | |  
 Db 34 GNICNDICV 43

## RESULT 13

O77634 PRELIMINARY; PRT; 155 AA.  
 ID O77634;  
 AC O77634;  
 DT 01-NOV-1998 (TrEMBLrel. 08, Created)  
 DT 01-NOV-1998 (TrEMBLrel. 08, Last sequence update)  
 DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)  
 DE Cellular disintegrin (Fragment).  
 GN ADAM-9.  
 OS Sus scrofa (Pig).

```

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
OX NCBI_TaxID=9823;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=99357011; PubMed=10429942;
RA Flannery C.R., Little C.B., Caterson B., Hughes C.E.;
RT "Effects of culture conditions and exposure to catabolic stimulators
RT (IL-1 and retinoic acid) on the expression of matrix
RT metalloproteinases (MMPs) and disintegrin metalloproteinases (ADAMs)
RT by articular cartilage chondrocytes.";
RL Matrix Biol. 18:225-237(1999).
DR EMBL; AF069646; AAC23530.1; -.
DR HSSP; P18619; 1FVL.
DR GO; GO:0007229; P: integrin-mediated signaling pathway; IEA.
DR InterPro; IPR006586; ADAM cysteine.
DR InterPro; IPR001762; Disintegrin.
DR Pfam; PF00200; disintegrin; 1.
DR PRINTS; PR00289; disintegrin.
DR ProDom; PD000664; Disintegrin; 1.
DR SMART; SM00608; ACR; 1.
DR SMART; SM00050; DISIN; 1.
DR PROSITE; PS0214; DISINTEGRIN_2; 1.
KW Integrin.
FT NON_TER 1
FT NON_TER 155
SQ SEQUENCE 155 AA; 16890 MW; 48A7CA6075F9390E CRC64;

Query Match 63.5%; Score 40; DB 6; Length 155;
Best Local Similarity 85.7%; Pred. No. 14;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GDVQCDC 7
DB 19 GDCCQDC 25

RESULT 14
QVFT8 ID Q7VFT8 PRELIMINARY; PRT; 181 AA.
AC Q7VFT8;
DT 01-OCT-2003 (TrEMBLrel. 25, Created)
DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Hypothetical protein.
GN HH1587.
OS Helicobacter hepaticus.
OC Bacteria; Proteobacteria; Epsilonproteobacteria; Campylobacterales;
OC Helicobacteraceae; Helicobacter.
OX NCBI_TaxID=32025;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 51449 / 3B1;
RX MEDLINE=22709201; PubMed=12810954;
RA Suerbaum S., Josenhans C., Sterzenbach T., Drescher B., Brandt P.,
RA Bell M., Droege M., Fartmann B., Fischer H.-P., Ge Z., Hoerster A.,
RA Holland R., Klein K., Koenig J., Macko L., Mendz G.L., Nyakatura G.,
RA Schauer D.B., Shen Z., Weber J., Frosch M., Fox J.G.;
RT "The complete genome sequence of the carcinogenic bacterium
RT Helicobacter hepaticus.";
RL Proc. Natl. Acad. Sci. U.S.A. 100:7901-7906(2003).
DR EMBL; AE017148; AAP78184.1; -.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 181 AA; 21202 MW; 4212C4E40E6CD486 CRC64;

Query Match 63.5%; Score 40; DB 16; Length 181;
Best Local Similarity 86.7%; Pred. No. 16;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 GDVQCDCIQ 9
DB 75 GDLIQDCVQ 83

RESULT 15
Q9ANV3 ID Q9ANV3 PRELIMINARY; PRT; 184 AA.
AC Q9ANV3;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE Putative ComF3 protein.
GN COMF3.
OS Lactobacillus delbrueckii.
OC Bacteria; Firmicutes; Lactobacillales; Lactobacillaceae;
OC Lactobacillus.
OX NCBI_TaxID=1584;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC11842;
RA van de Guchte M., Dervyn R., Ehrlich S.D., Maguin E.;
RT "L. bulgaricus ymdA - rf2 region.";
RL Submitted (NOV-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF320250; AAK00332.1; -.
SQ SEQUENCE 184 AA; 20908 MW; 6ECF46DAEC33C881 CRC64;

Query Match 63.5%; Score 40; DB 2; Length 184;
Best Local Similarity 71.4%; Pred. No. 16;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 GDVQCDC 7
DB 57 GEICQDC 63

Search completed: May 5, 2004, 13:33:42
Job time : 2.09687 secs

```

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 5, 2004, 13:16:48 ; Search time 0.271605 Seconds

(without alignments)  
2108.841 Million cell updates/sec

Title: US-09-743-684a-19

Perfect score: 63

Sequence: 1 GDVQCDCICQW 11

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 141681 seqs, 52070155 residues

Total number of hits satisfying chosen parameters: 141681

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt\_42:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	63	100.0	80	1	SAP_PIG
2	63	100.0	524	1	SAP_HUMAN
3	55	87.3	525	1	SAP_BOVIN
4	50	79.4	518	1	SAP_CHICK
5	47	74.6	554	1	SAP_RAT
6	47	74.6	557	1	SAP_MOUSE
7	39	61.9	512	1	VERV_THOXY
8	39	61.9	1696	1	PCRS_BRACL
9	38	60.3	355	1	CMG2_SCHPO
10	37	58.7	74	1	Y777_TREPA
11	37	58.7	374	1	NADA_HALNI
12	37	58.7	422	1	U183_HUMAN
13	37	58.7	422	1	U183_MOUSE
14	37	58.7	422	1	U183_RAT
15	37	58.7	748	1	CLPE_LACLA
16	37	58.7	748	1	CLPE_LACLC
17	37	58.7	1015	1	ITR4_TROME
18	37	58.7	5147	1	FAT_DROME
19	36	57.1	194	1	BC1B_HUMAN
20	36	57.1	330	1	EMB_MOUSE
21	36	57.1	425	1	CND0_HUMAN
22	36	57.1	425	1	CND0_MOUSE
23	36	57.1	511	1	GINB_PSEPL
24	36	57.1	564	1	TM16_HUMAN
25	36	57.1	757	1	COMP_HUMAN
26	36	57.1	819	1	AD09_HUMAN
27	35.5	56.3	1416	1	YN81_CAEBL
28	35	55.6	48	1	R332_MYCCE
29	35	55.6	48	1	R332_MYCPV
30	35	55.6	103	1	THB1_THORP
31	35	55.6	175	1	CEN_ARAIH
32	35	55.6	181	1	Y4AS_RHISN
33	35	55.6	199	1	AA27_MOUSE

34	35	55.6	228	1	VIF_OMVVS	P16902 ovine lenti
35	35	55.6	230	1	VIF_VILV	P03403 viana lenti
36	35	55.6	230	1	VIF_VILV1	P23430 viana lenti
37	35	55.6	285	1	TI3B_HUMAN	O97275 homo sapien
38	35	55.6	309	1	TI3B_MOUSE	O97472 mus musculu
39	35	55.6	477	1	TM17_MOUSE	O76PM3 mus musculu
40	35	55.6	477	1	TM17_RAT	O9WV59 rattus norv
41	35	55.6	498	1	VAC3_SCHPO	O09817 schizosacch
42	35	55.6	752	1	CLPE_STRPN	P35594 streptococc
43	35	55.6	867	1	ATX7_MOUSE	O87411 mus musculu
44	35	55.6	892	1	ATX7_HUMAN	O15265 homo sapien
45	35	55.6	961	1	TSP4_HUMAN	P35443 homo sapien

## ALIGNMENTS

## RESULT 1

ID	SAP_PIG	STANDARD	PRT	80 AA.
AC	P81405			
DT	15-DEC-1998 (Rel. 37, Created)			
DT	16-OCT-2001 (Rel. 40, Last sequence update)			
DT	28-FEB-2003 (Rel. 41, Last annotation update)			
DE	Saposin B (Cerebroside sulfate activator) (CS-ACT) (Non-specific activator) (Sphingolipid activator protein 1) (SAP-1).			
OS	Sus scrofa (pig)			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OX	Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.			
NCBI_TaxID=9823;				
RN	[1]			
RP	SEQUENCE OF 1-79.			
RC	TISSUE=Kidney;			
RA	MEDLINE=93229306; PubMed=8471613;			
RA	Stevens R.L., Fauli K.F., Conklin K.A., Green B.N., Fluharty A.L.;			
RT	"porcine cerebroside sulfate activator: further structural characterization and disulfide identification.";			
RL	Biochemistry 32:4051-4059(1993).			
RP	[2]			
RC	SEQUENCE OF 1-64.			
RA	TISSUE=Kidney;			
RA	MEDLINE=92222651; PubMed=1562358;			
RT	Fluharty A.L., Katona Z., Meek W.B., Frei K., Fowler A.V.;			
RT	"The cerebroside sulfate activator from pig kidney: purification and molecular structure.";			
RL	Biochem. Med. Metab. Biol. 47:66-85(1992).			
RP	[3]			
RC	STRUCTURE OF CARBOHYDRATE ON ASN-21.			
RA	MEDLINE=21110404; PubMed=11180632;			
RA	Fauli K.F., Johnson J., Kim M.J., To T., Whitelegge J.P.,			
RT	Stevens R.L., Fluharty C.B., Fluharty A.L.;			
RT	"Structure of the asparagine-linked sugar chains of porcine kidney and human urine cerebroside sulfate activator protein.";			
RL	J. Mass Spectrom. 35:1416-1424(2000).			
RP	[4]			
RC	MASS SPECTROMETRY.			
RA	TISSUE=Kidney;			
RA	MEDLINE=99441404; PubMed=10510427;			
RA	Fauli K.F., Whitelegge J.P., Higginson J., To T., Johnson J.,			
RA	Krutchinsky A.N., Standing K.G., Waring A.J., Stevens R.L.,			
RT	Fluharty C.B., Fluharty A.L.;			
RT	"Cerebroside sulfate activator protein (Saposin B): chromatographic and electrospray mass spectrometric properties.";			
RL	J. Mass Spectrom. 34:1040-1054(1999).			
RP	[1]			
RC	FUNCTION: Saposin B stimulates the hydrolysis of galacto-			
CC	cerebroside sulfate by arylsulfatase A (EC 3.1.6.8), GMI			
CC	galactosylceramide by beta-galactosidase (EC 3.2.1.23) and			
CC	globosylceramide by alpha-galactosidase A (EC 3.2.1.22).			
CC	Saposin B forms a solubilizing complex with the substrates of the			
CC	sphingolipid hydrolases.			
CC	SUBUNIT: Saposin B is a homodimer (by similarity).			
CC	-1- PTM: The one residue extended Saposin B-Val is only found in a			
CC	minority of the chains.			

CC -1- SIMILARITY: Contains 1 saposin B-type domain.  
 DR GlycosultideB; P81405; -;  
 DR InterPro; IPR007856; SApB\_1.  
 DR InterPro; IPR008138; SApB\_2.  
 DR InterPro; IPR008373; Saposin.  
 DR InterPro; IPR008139; SaposinB.  
 DR Pfam; PF05184; SApB\_1; 1.  
 DR Pfam; PF03489; SApB\_2; 1.  
 DR PRINTS; PR01797; SAPOSIN.  
 DR SMART; SM00119; SApB; 1.  
 KW Glycoprotein; sphingolipid metabolism.  
 FT CHAIN 1  
 FT DOMAIN 1 80  
 FT DISULFID 4 77  
 FT DISULFID 7 71  
 FT DISULFID 36 47  
 FT CARBOHYD 21 21  
 SQ SEQUENCE 80 AA; 8949 MW; EF7BA249B63E789C CRC64;  
 /FTID=CAR\_000177.  
 N-LINKED (GLCNAC. . .) (COMPLEX).  
 Query Match 100.0%; Score 63; DB 1; Length 80;  
 Best Local Similarity 100.0%; Pred. No. 0.00038;  
 Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 GDVQCDCIQMV 11  
 DB 1 GDVQCDCIQMV 11  
 RESULT 2  
 SAp\_HUMAN  
 ID\_SAP\_HUMAN STANDARD; PRT; 524 AA.  
 AC P07602; P07292; P15793; P78538; P78541; P78546; P78547; P78558;  
 AC Q92739; Q92740; Q92741; Q92742;  
 DT 01-APR-1988 (Rel. 07; Created)  
 DT 01-APR-1990 (Rel. 14; Last sequence update)  
 DT 10-OCT-2003 (Rel. 42; Last annotation update)  
 DE Proactivator polypeptide precursor [Contains: Saposin A (Protein A);  
 DE Saposin B (Sphingolipid activator protein 1) (SAP-1) (Cerebroside  
 DE sulfate activator) (CSAct) (Dispersin) (Sulfatide/GM1 activator);  
 DE Saposin C (Co-beta-glucosidase) (A1 activator) (Glucosylceramidase  
 DE activator) (Sphingolipid activator protein 2) (SAP-2); Saposin D  
 DE (Protein C) (Component C)].  
 GN PSAP.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Liver;  
 RX MEDLINE=90129043; PubMed=2515150;  
 RA Roman E.G., Grabowski G.A.;  
 RT "Molecular cloning of a human co-beta-glucosidase cDNA: evidence that  
 RT four sphingolipid hydrolase activator proteins are encoded by single  
 RT genes in humans and rats.";  
 RL Genomics 5:486-492(1989).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Liver;  
 RX MEDLINE=9255151; PubMed=2498298;  
 RA Nakano T., Sandhoff K., Stuenkel J., Christomanou H., Suzuki K.;  
 RT "Structure of full-length cDNA coding for sulfatide activator, a  
 RT co-beta-glucosidase and two other homologous proteins: two alternate  
 RT forms of the sulfatide activator.";  
 RL J. Biochem. 105:152-154(1989).  
 RN [3]  
 RP SEQUENCE FROM N.A. (ISOFORM SAP-WU-0).  
 RC TISSUE=Brain, Eye, and Skin;  
 RX MEDLINE=22388257; PubMed=12477932;  
 RA Strauberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,  
 RA Klausner R.D., Collins F.S., Wagner L., Shennan C.M., Schuler G.D.,  
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,

RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Haieh F.,  
 RA Diachenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,  
 RA Stapleton M., Soares M.B., Bonaldi M.F., Casavant T.L., Scheetz T.E.,  
 RA Brownstein M.J., Uedlin T.B., Toshiyuki S., Carninci P., Prange C.,  
 RA Rahn S.S., Locuelli N.A., Peters G.J., Abramson R.D., Mullaly S.J.,  
 RA Boask S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,  
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,  
 RA Villalon D.K., Wuzny D.W., Sodergren E.J., Lu X., Gibbs R.A.,  
 RA Fahy J., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A.,  
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,  
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,  
 RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smalhus D.E.,  
 RA Scherzer A., Schein J.E., Jones S.J.W., Marra M.A.;  
 RT "Generation and initial analysis of more than 15,000 full-length  
 RT human and mouse cDNA sequences.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).  
 RN [4]  
 RP SEQUENCE OF 59-125 AND 304-513 FROM N.A.  
 RC TISSUE=Brain;  
 RX MEDLINE=91192146; PubMed=2013321;  
 RA Holtschmidt H., Sandhoff K., Fuerst W., Kwon H.Y., Schnabel D.,  
 RA Suzuki K.;  
 RT "The organization of the gene for the human cerebroside sulfate  
 RT activator protein.";  
 RL FEBS Lett. 280:267-270(1991).  
 RN [5]  
 RP SEQUENCE OF 164-524 FROM N.A.  
 RX MEDLINE=88068647; PubMed=2825202;  
 RA Dewi N.N., Wenger D.A., O'Brien J.S.;  
 RT "Nucleotide sequence of cloned cDNA for human sphingolipid activator  
 RT protein 1 precursor.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 84:8652-8656(1987).  
 RN [6]  
 RP PARTIAL SEQUENCE OF 60-142.  
 RX MEDLINE=9240739; PubMed=2717620;  
 RA Morimoto S., Martin B.M., Yamamoto Y., Kretz K.A., O'Brien J.S.,  
 RA Kishimoto Y.;  
 RT "Saposin A: second cerebroside activator protein.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 86:3389-3393(1989).  
 RN [7]  
 RP SEQUENCE OF 195-263 FROM N.A.  
 RX MEDLINE=86130593; PubMed=2868718;  
 RA Dewi N.N., Wenger D.A., Fujibayashi S., Donoviel M., Esch F.,  
 RA Hill F., O'Brien J.S.;  
 RT "Molecular cloning of the sphingolipid activator protein-1 (SAP-1),  
 RT the sulfate sulfatase activator.";  
 RL Biochem. Biophys. Res. Commun. 134:989-994(1986).  
 RN [8]  
 RP SEQUENCE OF 195-274.  
 RC TISSUE=Kidney;  
 RX MEDLINE=91006165; PubMed=2209618;  
 RA First W., Schubert J., Machleidt W., Meyer H.E., Sandhoff K.;  
 RT "The complete amino-acid sequences of human ganglioside GM2 activator  
 RT protein and cerebroside sulfate activator protein.";  
 RL Eur. J. Biochem. 192:709-714(1990).  
 RN [9]  
 RP SEQUENCE OF 195-274.  
 RX MEDLINE=89207118; PubMed=3242555;  
 RA Kleinschmidt T., Christomanou H., Braunitzer G.;  
 RT "Complete amino-acid sequence of the naturally occurring A2 activator  
 RT protein for enzymic sphingomyelin degradation: identity to the  
 RT sulfate activator protein (SAP-1).";  
 RL Biol. Chem. Hoppe-Seyler 369:1361-1365(1988).  
 RN [10]  
 RP SEQUENCE OF 311-390.  
 RX MEDLINE=88163077; PubMed=3442600;  
 RA Kleinschmidt T., Christomanou H., Braunitzer G.;  
 RT "Complete amino-acid sequence and carbohydrate content of the  
 RT naturally occurring glucosylceramide activator protein (A1 activator)  
 RT absent from a new human Gaucher disease variant.";  
 RL Biol. Chem. Hoppe-Seyler 368:1571-1578(1987).  
 RN [11]

RP SEQUENCE OF 407-484.  
 RX MEDLINE=89000190; PubMed=3048308;  
 RA Furst W., Machleidt W., Sandhoff K.;  
 RT "The precursor of sulfatide activator protein is processed to three  
 RT different proteins.";  
 RL Biol. Chem. Hoppe-Seyler 369:317-328(1988).  
 RN [12]  
 RP PARTIAL SEQUENCE OF 405-484.  
 RX MEDLINE=89025876; PubMed=2845979;  
 RA Morimoto S., Martin B.M., Kishimoto Y., O'Brien J.S.;  
 RT "Saposin D: a sphingomyelinase activator.";  
 RL Biochem. Biophys. Res. Commun. 156:403-410(1988).  
 RN [13]  
 RP SEQUENCE OF 17-26.  
 RC TISSUE=Milk;  
 RX MEDLINE=92068206; PubMed=1958198;  
 RA Kondoh K., Haheno T., Sano A., Kakimoto Y.;  
 RT "Isolation and characterization of prosaposin from human milk.";  
 RL Biochem. Biophys. Res. Commun. 181:286-292(1991).  
 RN [14]  
 RP PARTIAL SEQUENCE (SAPOSIN B), AND STRUCTURE OF CARBOHYDRATES.  
 RC TISSUE=Urine;  
 RX MEDLINE=20032116; PubMed=10562467;  
 RA Fluharty A.L., Lombardo C., Louis A., Stevens R.L., Whitelegge J.P.,  
 RA Waring A.J., To T., Fluharty C.B., Faull K.F.;  
 RT "Preparation of the cerebroside sulfate activator (CSact or saposin B)  
 RT from human urine.";  
 RL Mol. Genet. Metab. 68:391-403(1999).  
 RN [15]  
 RP STRUCTURE OF CARBOHYDRATE ON ASN-215.  
 RX MEDLINE=21110404; PubMed=1180632;  
 RA Faull K.F., Johnson J., Kim M.J., To T., Whitelegge J.P.,  
 RA Stevens R.L., Fluharty C.B., Fluharty A.L.;  
 RT "Structure of the asparagine-linked sugar chains of porcine kidney and  
 RT human urine cerebroside sulfate activator protein.";  
 RL J. Mass Spectrom. 35:1416-1424(2000).  
 RN [16]  
 RP SAPOSIN D DISULFIDE BONDS.  
 RX MEDLINE=99337688; PubMed=10406958;  
 RA Tatti M., Salvioli R., Clafioni F., Pucci P., Andolfo A.,  
 RA Amoresano A., Vaccaro A.M.;  
 RT "Structural and membrane-binding properties of saposin D.";  
 RL Eur. J. Biochem. 263:486-494(1999).  
 RN [17]  
 RP SAPOSIN B DISULFIDE BONDS.  
 RX MEDLINE=22398398; PubMed=12510003;  
 RA Ahn V.E., Faull K.F., Whitelegge J.P., Higginson J., Fluharty A.L.,  
 RA Prive G.G.;  
 RT "Expression, purification, crystallization, and preliminary X-ray  
 RT analysis of recombinant human saposin B.";  
 RL Protein Expr. Purif. 27:186-193(2003).  
 RN [18]  
 RP MASS SPECTROMETRY.  
 RC TISSUE=Urine;  
 RX MEDLINE=99441404; PubMed=10510427;  
 RA Faull K.F., Whitelegge J.P., Higginson J., To T., Johnson J.,  
 RA Krutcheney A.N., Standing K.G., Waring A.J., Stevens R.L.,  
 RA Fluharty C.B., Fluharty A.L.;  
 RT "Cerebroside sulfate activator protein (Saposin B): chromatographic  
 RT and electrospray mass spectrometric properties.";  
 RL J. Mass Spectrom. 34:1040-1054(1999).  
 RN [19]  
 RP X-RAY CRYSTALLOGRAPHY (2.2 ANGSTROMS) OF 195-273, AND MUTAGENESIS OF  
 RP ILE-240.  
 RX MEDLINE=22406333; PubMed=12518053;  
 RA Ahn V.E., Faull K.F., Whitelegge J.P., Fluharty A.L., Prive G.G.;  
 RT "Crystal structure of saposin B reveals a dimeric shell for lipid  
 RT binding.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 100:38-43(2003).  
 RN [20]  
 RP REVIEW ON MLD VARIANTS.  
 RX MEDLINE=95170731; PubMed=7866401;  
 RA Gieselmann V., Zlotogora J., Harris A., Wenger D.A., Morris C.P.;

RT "Molecular genetics of metachromatic leukodystrophy.";  
 RL Hum. Mutat. 4:233-242(1994).  
 RN [21]  
 RP VARIANT MLD ILE-217.  
 RX MEDLINE=90147748; PubMed=2302219;  
 RA Rafi M.A., Zhang X.-L., Degala G., Wenger D.A.;  
 RT "Detection of a point mutation in sphingolipid activator protein-1  
 RT mRNA in patients with a variant form of metachromatic  
 RT leukodystrophy.";  
 RL Biochem. Biophys. Res. Commun. 166:1017-1023(1990).  
 RN [22]  
 RP SEQUENCE FROM N.A., AND VARIANT MLD ILE-217.  
 RX MEDLINE=90207231; PubMed=2320574;  
 RN [23]  
 Query Match 100.0%; Score 63; DB 1; Length 524;  
 Best Local Similarity 100.0%; Pred. No. 0.0023;  
 Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 Oy 1 GDVCGDCTQWV 11  
 Db 195 GDVCGDCTQWV 205  
 RESULT 3  
 SLP\_BOVIN STANDARD; PRT; 525 AA.  
 AC P26779; Q9N2G4;  
 DT 01-AUG-1992 (Rel. 23, Created)  
 DT 16-OCT-2001 (Rel. 40, Last sequence update)  
 DT 16-OCT-2001 (Rel. 40, Last annotation update)  
 DE Proactivator polypeptide precursor [contains: Saposin A (Protein A);  
 DE Saposin B (Sphingolipid activator protein 1) (SAP-1) (Cerebroside  
 DE sulfate activator) (CSact) (Dispersin) (Sulfatide/GM1 activator);  
 DE Saposin C (Co-beta-glucosidase) (AI activator) (Glucosylceramidase  
 DE activator) (Sphingolipid activator protein 2) (SAP-2); Saposin D  
 DE (Protein C) (Component C)].  
 GN PSAP.  
 OS Bos taurus (Bovine).  
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;  
 OC Bovidae; Bovinae; Bos.  
 OX NCBI\_TaxID=9913;  
 RN [1]  
 RP SEQUENCE FROM N.A., AND VARIANTS.  
 RC TISSUE=Mammary gland;  
 RA Azuma N., Yoshida K.;  
 RT "RT-PCR cloning of bovine prosaposin.";  
 RL Submitted (JAN-2000) to the EMBL/Genbank/DBJ databases.  
 RN [2]  
 RP SEQUENCE OF 312-391.  
 RC TISSUE=Spleen;  
 RX MEDLINE=92207994; PubMed=1554743;  
 RA Sano A., Mizuno T., Kondoh K., Haheno T., Ueno S.-I., Kakimoto Y.,  
 RA Morita N.;  
 RT "Saposin-C from bovine spleen: complete amino acid sequence and  
 RT relation between the structure and its biological activity.";  
 RL Biochim. Biophys. Acta 1120:75-80(1992).  
 CC -1- FUNCTION: The lysosomal degradation of sphingolipids takes place  
 CC by the sequential action of specific hydrolases. Some of these  
 CC enzymes require specific low-molecular mass, non-enzymic proteins:  
 CC the sphingolipids activator proteins (coproteins) (By similarity).  
 CC -1- FUNCTION: Saposin A and saposin C stimulate the hydrolysis of  
 CC glucosylceramide by beta-glucosylceramidase (EC 3.2.1.45) and  
 CC galactosylceramide by beta-galactosylceramidase (EC 3.2.1.46).  
 CC Saposin C apparently acts by combining with the enzyme and acidic  
 CC lipid to form an activated complex, rather than by solubilizing  
 CC the substrate.  
 CC -1- FUNCTION: Saposin B stimulates the hydrolysis of galacto-  
 CC cerebroside sulfate by arylsulfatase A (EC 3.1.6.8), GM1  
 CC gangliosides by beta-galactosidase (EC 3.2.1.23) and  
 CC globotriaosylceramide by alpha-galactosidase A (EC 3.2.1.22).  
 CC Saposin B forms a solubilizing complex with the substrates of the  
 CC sphingolipid hydrolases (By similarity).

CC	- FUNCTION: Saposin D is a specific sphingomyelin phosphodiesterase
CC	activator (EC 3.1.4.12) (By similarity).
CC	- SUBUNIT: Saposin B is a homodimer (By similarity).
CC	- SUBCELLULAR LOCATION: Lysosomal.
CC	- PTM: This precursor is proteolytically processed to 4 small
CC	peptides, which are similar to each other and are sphingolipid
CC	hydrolase activator proteins (By similarity).
CC	- SIMILARITY: Contains 2 saposin A-type domains.
CC	- SIMILARITY: Contains 4 saposin B-type domains.
CC	-----
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC	the European Bioinformatics Institute. There are no restrictions on its
CC	use by non-profit institutions as long as its content is in no way
CC	modified and this statement is not removed. Usage by and for commercial
CC	entities requires a license agreement (See <a href="http://www.isdb.ch/announce/">http://www.isdb.ch/announce/</a>
CC	or send an email to <a href="mailto:license@isdb-sib.ch">license@isdb-sib.ch</a> ).
CC	-----
DR	EMBL; AB036791; BA95677.1; -.
DR	InterPro; IPR003119; SAPA_1.
DR	InterPro; IPR007856; SapB_1.
DR	InterPro; IPR008138; SAPB_1.
DR	InterPro; IPR008140; SapB_sub.
DR	InterPro; IPR008373; Saposin.
DR	InterPro; IPR008139; SaposinB.
DR	Pfam; PF02199; SAPA; 2.
DR	Pfam; PF03489; SapB_1; 4.
DR	Pfam; PF03489; SapB_2; 4.
DR	PRINTS; PR01797; SAFOSIN.
DR	ProDom; PD001732; SapB_sub; 3.
DR	SMART; SM00162; SAPA_2.
DR	SMART; SM00118; SAPB; 4.
KW	Signal; Glycoprotein; Lysosome; Sphingolipid metabolism; Repeat.
FT	SIGNAL 1 16 POTENTIAL.
FT	PROPEP 17 58
FT	CHAIN 60 142 SAPOSIN A.
FT	PROPEP 144 195
FT	CHAIN 196 275 SAPOSIN B.
FT	PROPEP 277 310
FT	CHAIN 312 392 SAPOSIN C.
FT	PROPEP 393 404
FT	CHAIN 406 487
FT	PROPEP 489 525 SAPOSIN D.
FT	DOMAIN 21 54
FT	DOMAIN 59 142 SAPOSIN-LIKE TYPE A 1.
FT	DOMAIN 194 276 SAPOSIN-LIKE TYPE B 1.
FT	DOMAIN 312 393 SAPOSIN-LIKE TYPE B 2.
FT	DOMAIN 406 487 SAPOSIN-LIKE TYPE B 3.
FT	DOMAIN 492 525 SAPOSIN-LIKE TYPE B 4.
FT	DISULFID 63 138 BY SIMILARITY.
FT	DISULFID 66 132 BY SIMILARITY.
FT	DISULFID 94 106 BY SIMILARITY.
FT	DISULFID 199 272 BY SIMILARITY.
FT	DISULFID 202 266 BY SIMILARITY.
FT	DISULFID 231 242 BY SIMILARITY.
FT	DISULFID 316 389 BY SIMILARITY.
FT	DISULFID 319 383 BY SIMILARITY.
FT	DISULFID 347 358 BY SIMILARITY.
FT	DISULFID 410 483 BY SIMILARITY.
FT	DISULFID 413 477 BY SIMILARITY.
FT	DISULFID 441 452 BY SIMILARITY.
FT	CARBOHYD 80 N-LINKED (GLCNAC. . . ) (POTENTIAL).
FT	CARBOHYD 101 N-LINKED (GLCNAC. . . ) (POTENTIAL).
FT	CARBOHYD 216 N-LINKED (GLCNAC. . . ) (POTENTIAL).
FT	CARBOHYD 333 N-LINKED (GLCNAC. . . ) (POTENTIAL).
FT	CARBOHYD 427 N-LINKED (GLCNAC. . . ) (POTENTIAL).
FT	VARIANT 127 263 M-Q-P->IIR.
FT	VARIANT 127 263 M-Q-P->Q (IN REF. 2).
FT	CONFLICT 317 R->S (IN REF. 1).
FT	CONFLICT 367 R->S (IN REF. 1).
SO	SEQUENCE 525 AA; 58120 MW; 293AACDFB9C4FA99 CRC64;

[illegible]

RA	Blaas M.W., Sylvester S.R., Tsuruta J.K., Griswold M.D.; "Biosynthesis and molecular cloning of sulfated glycoprotein 1 secreted by rat Sertoli cells: sequence similarity with the 70-kilodalton precursor to sulfate/GM1 activator"; <i>Biochemistry</i> 27:4557-4564(1988).
RN	[2]
RP	SEQUENCE FROM N.A.
RC	TISSUE=Testis;
RX	MEDLINE=96128541; PubMed=8573994; Morales C.R., El-Alfy M., Zhao Q., Igodura S.A.;
RA	"Molecular role of sulfated glycoprotein-1 (SGP-1/prosaposin) in Sertoli cells."; <i>Histol. Histopathol.</i> 10:1023-1034(1995).
RL	[3]
RN	SEQUENCE FROM N.A.
RP	TISSUE=Testicle;
RC	MEDLINE=96175245; PubMed=8601692; Morales C.R., El-Alfy M., Zhao Q., Igodura S.A.;
RX	"Expression and tissue distribution of rat sulfated glycoprotein-1 (prosaposin)."; <i>J. Histochem. Cytochem.</i> 44:327-337(1996).
RT	-I- SUBCELLULAR LOCATION: Extracellular. -I- SIMILARITY: Contains 2 saposin A-type domains. -I- SIMILARITY: Contains 4 saposin B-type domains.
CC	-----
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outbreak - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial/ entities requires a license agreement (See <a href="http://www.isb-sib.ch/announce/">http://www.isb-sib.ch/announce/</a> or send an email to <a href="mailto:license@isb-sib.ch">license@isb-sib.ch</a> ). -----
CC	EMBL; M19936; AAA42136.1; -; DR EMML; S81353; AAB36042.2; -; DR EMBL; S81373; AAB36233.2; -; PIR; A28716; A28716.
DR	InterPro; IPR003119; Sapa. DR InterPro; IPR007856; SApB_1. DR InterPro; IPR008138; SApB_2. DR InterPro; IPR008140; SApB_sub. DR InterPro; IPR008373; Saposin. DR InterPro; IPR008139; SaposinB.
DR	Pfam; PF02199; SAPA; 2. DR Pfam; PF05184; SApB_1; 3. DR Pfam; PF03489; SApB_2; 4. DR PRINTS; PRO1797; SAPOSIN. DR ProDom; PD001732; SApB_sub; 3.
DR	SMART; SM00162; SAPA; 2. DR SMART; SM00118; SAPB; 4.
KW	SIGNAL; Glycoprotein; Repeat.
FT	SIGNAL 1 16
FT	CHAIN 17 554 SUPDATED GLYCOPROTEIN 1.
FT	DOMAIN 21 54
FT	DOMAIN 21 138 SAPOSIN-LIKE TYPE A 1.
FT	DOMAIN 193 274 SAPOSIN-LIKE TYPE B 1.
FT	DOMAIN 310 391 SAPOSIN-LIKE TYPE B 2.
FT	DOMAIN 435 516 SAPOSIN-LIKE TYPE B 3.
FT	DOMAIN 521 554 SAPOSIN-LIKE TYPE B 4.
FT	DISULFID 63 138 BY SIMILARITY.
FT	DISULFID 66 132 BY SIMILARITY.
FT	DISULFID 94 106 BY SIMILARITY.
FT	DISULFID 197 270 BY SIMILARITY.
FT	DISULFID 200 264 BY SIMILARITY.
FT	DISULFID 229 240 BY SIMILARITY.
FT	DISULFID 314 387 BY SIMILARITY.
FT	DISULFID 317 381 BY SIMILARITY.
FT	DISULFID 345 356 BY SIMILARITY.
FT	DISULFID 439 512 BY SIMILARITY.
FT	DISULFID 442 506 BY SIMILARITY.
FT	DISULFID 470 481 BY SIMILARITY.
FT	CARBHYD 80 80 N-LINKED (GLCNAC... ) (POTENTIAL).
FT	CARBHYD 214 214 N-LINKED (GLCNAC... ) (POTENTIAL).

FT CARBOHYD 331 331 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 456 456 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CONFLICT 115 115 P -> L (IN REF. 2 AND 3).  
 FT CONFLICT 299 299 D -> R (IN REF. 2).  
 FT CONFLICT 462 462 I -> V (IN REF. 3).  
 FT CONFLICT 527 527 W -> R (IN REF. 3).  
 FT CONFLICT 536 536 S -> M (IN REF. 3).  
 SQ SEQUENCE 554 AA; 61123 MW; DFE3F3A3A0520C6B CRC64;

Query Match 74.6%; Score 47; DB 1; Length 554;  
 Best Local Similarity 70.0%; Pred. No. 1.1;  
 Matches 7; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 2 DVCQDCIQMV 11  
 DB 195 DVCQDCIKLV 204

RESULT 6  
 SLP\_MOUSE STANDARD; PRT; 557 AA.  
 AC 061207; 060861; 064006; 064219;  
 DT 15-DEC-1998 (Rel. 37, Created)  
 DT 15-DEC-1998 (Rel. 37, Last sequence update)  
 DT 15-MAR-2004 (Rel. 43, Last annotation update)  
 DE Sulfated glycoprotein 1 precursor (SGP-1) (Prosaposin).  
 GN PSAP OR SGP.  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 NC NCB1; TaxID=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Liver;  
 RA Tsuda M., Sakiyama T., Endo H., Kitagawa T.;  
 RT "The primary structure of mouse saposin.";  
 RL Biochem. Biophys. Res. Commun. 184:1266-1272(1992).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC MEDLINE=94272317; PubMed=8003952;  
 RA Sprecher-Levy H., Orr-Urtreger A., Lonai P., Horowitz M.;  
 RT "Murine prosaposin: expression in the reproductive system of a gene  
 RL Cell. Mol. Biol. 40:233-233(1994).  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RC MEDLINE=96084310; PubMed=8565332;  
 RA Cao Q.P., Crain W.R.;  
 RT "Expression of SGP-1 mRNA in preimplantation mouse embryos.";  
 RL Dev. Genet. 17:263-271(1995).  
 RN [4]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=BALB/c; TISSUE=Liver;  
 RA Zhao Q.Q., Hay N.N., Morales C.R.;  
 RT Submitted (MAY-1996) to the EMBL/GenBank/DBJ databases.  
 RL -1- SUBCELLULAR LOCATION: Extracellular.  
 CC -1- SIMILARITY: Contains 2 saposin A-type domains.  
 CC -1- SIMILARITY: Contains 4 saposin B-type domains.  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.1sb-sib.ch/announce/>  
 CC or send an email to [license@1sb-sib.ch](mailto:license@1sb-sib.ch)).  
 CC -----  
 CC EMBL; S36200; AAB22175.1; -;  
 DR EMBL; S71616; AAB31059.1; -;  
 DR EMBL; U27340; AAA92567.1; -;  
 DR EMBL; U57999; AAB02695.1; -;  
 DR PIR; JH0604; JH0604.

DR MGD; MGI:97783; Psap.  
 DR InterPro; IPR003119; Sapa.  
 DR InterPro; IPR007856; Sapa\_1.  
 DR InterPro; IPR008138; Sapa\_2.  
 DR InterPro; IPR008140; Sapa\_sub.  
 DR InterPro; IPR008373; Saposin.  
 DR InterPro; IPR008339; SaposinB.  
 DR Pfam; PF02199; Sapa\_2.  
 DR Pfam; PF05184; Sapa\_1; 4.  
 DR Pfam; PF03489; Sapa\_2; 4.  
 DR PRINTS; PR01797; Saposin.  
 DR ProDom; PD001732; Sapa\_sub; 3.  
 DR SMART; SM00152; Sapa; 7.  
 DR SMART; SM00118; Sapa; 4.  
 KM Signal; Glycoprotein; Repeat.  
 FT SIGNAL 1 16  
 FT CHAIN 17 557  
 FT DOMAIN 21 54  
 FT DOMAIN 59 142  
 FT DOMAIN 193 277  
 FT DOMAIN 313 394  
 FT DOMAIN 438 519  
 FT DOMAIN 524 557  
 FT DISULFID 63 138  
 FT DISULFID 66 132  
 FT DISULFID 94 106  
 FT DISULFID 197 273  
 FT DISULFID 200 267  
 FT DISULFID 229 240  
 FT DISULFID 317 390  
 FT DISULFID 320 384  
 FT DISULFID 348 359  
 FT DISULFID 442 515  
 FT DISULFID 445 509  
 FT DISULFID 473 484  
 FT CARBOHYD 80 80  
 FT CARBOHYD 214 214  
 FT CARBOHYD 334 334  
 FT CARBOHYD 459 459  
 FT CONFLICT 83 83  
 FT CONFLICT 158 158  
 FT CONFLICT 160 160  
 FT CONFLICT 171 172  
 FT CONFLICT 244 244  
 FT CONFLICT 254 254  
 FT CONFLICT 255 255  
 FT CONFLICT 260 260  
 FT CONFLICT 307 307  
 FT CONFLICT 322 322  
 FT CONFLICT 349 350  
 FT CONFLICT 367 367  
 FT CONFLICT 370 370  
 FT CONFLICT 373 373  
 FT CONFLICT 391 391  
 FT CONFLICT 393 393  
 FT CONFLICT 406 406  
 FT CONFLICT 430 430  
 FT CONFLICT 445 445  
 FT CONFLICT 448 448  
 SQ SEQUENCE 557 AA; 61422 MW; 134593E20499E35E CRC64;

Query Match 74.6%; Score 47; DB 1; Length 557;  
 Best Local Similarity 70.0%; Pred. No. 1.1;  
 Matches 7; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 2 DVCQDCIQMV 11  
 DB 195 DVCQDCIKLV 204

RESULT 7  
 VENV\_THOQV STANDARD; PRT; 512 AA.  
 ID VENV\_THOQV



```

AC P286977
BT 01-DEC-1992 (Rel. 24, Created)
DT 01-DEC-1992 (Rel. 24, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE Envelope glycoprotein precursor (Surface glycoprotein 75) .
GN P4.
OS Thogoto virus (isolate Siar 126) (Tho) .
OC Viruses; ssRNA negative-strand viruses; Orthomyxoviridae;
OC Thogotovirus.
XX NCBI_TaxID=126796;
RN (1)
RP SEQUENCE FROM N.A.
RX MEDLINE=92124738; PubMed=1733105;
RA Morse M.A., Marriott A.C., Nuttall P.A.;
RT "The glycoprotein of Thogoto virus (a tick-borne orthomyxo-like
RT virus) is related to the baculovirus glycoprotein GP64." ;
RL Virology 186:640-646(1992) .
CC -1- FUNCTION: POSSIBLE ROLE IN ENDOCYTIC FUSION EVENTS DURING
CC INFECTION.
CC -1- SUBUNIT: Monomer (Probable) .
CC -1- SIMILARITY: TO DHORI VIRUS ENVELOPE GLYCOPROTEIN AND TO
CC BACULOVIRUSES MAJOR ENVELOPE GLYCOPROTEIN (P64/P67) .
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch) .
CC -----
CC
CC EMBL; M77280; AAA47918.1; - ;
CC
CC PIR; A40821; VGIYTH.
CC
CC InterPro; IPR004955; Baculo gp64.
CC
CC Pfam; PF03273; Baculo gp64_1
CC
CC Glycoprotein; Transmembrane; Signal.
CC
CC FT SIGNAL 1 15 POTENTIAL.
CC FT CHAIN 16 512 ENVELOPE GLYCOPROTEIN.
CC FT TRANSMEM 479 502 POTENTIAL.
CC FT CARBOHYD 185 185 N-LINKED (GLCNAC . . .) (POTENTIAL) .
CC FT CARBOHYD 263 263 N-LINKED (GLCNAC . . .) (POTENTIAL) .
CC FT CARBOHYD 289 289 N-LINKED (GLCNAC . . .) (POTENTIAL) .
CC FT CARBOHYD 378 378 N-LINKED (GLCNAC . . .) (POTENTIAL) .
CC FT CARBOHYD 416 416 N-LINKED (GLCNAC . . .) (POTENTIAL) .
CC
CC SQ SEQUENCE 512 AA; 57550 MW; 0398FC36284ADDF1 CRC64;
CC
Query March 61.9%; Score 39; DB 1; Length 512;
Best Local Similarity 71.4%; Pred. No. 23;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0.
OY 1 GDVCDGC 7
Db 109 GDICDGC 115

```

```

RT "Evolution of the prohormone convertases: identification of a
RL homologue of PC6 in the protochordate amphioxus.",
CC Biochim. Biophys. Acta 1477:338-348(2000).
CC -1- FUNCTION: Likely to represent a widespread endoprotease activity
CC within the constitutive and regulated secretory pathway. Capable
CC of cleavage at the RX(K/R)R consensus motif (By similarity).
CC -1- CATALYTIC ACTIVITY: Release of mature proteins from their
CC proproteins by cleavage of Arg-Xaa-Yaa-Arg-|-Zaa bonds, where Xaa
CC can be any amino acid and yaa is Arg or Lys.
CC -1- SUBCELLULAR LOCATION: ISOFORM A AND ISOFORM C ARE SECRETED.
CC ISOFORM B IS A TYPE I MEMBRANE PROTEIN.
CC -1- ALTERNATIVE PRODUCTS:
CC Event=Alternative splicing; Named isoforms=3;
CC Name=B;
CC IsoId=Q9NUJ5-1; Sequence=Displayed;
CC Name=A;
CC IsoId=Q9NUJ5-2; Sequence=VSP_005444, VSP_005445;
CC Name=C;
CC IsoId=Q9NUJ5-3; Sequence=VSP_005442, VSP_005443;
CC -1- DOMAIN: The propeptide domain acts as an intramolecular chaperone
CC assisting the folding of the zymogen within the endoplasmic
CC reticulum.
CC -1- SIMILARITY: Belongs to peptidase family S8.
CC -1- SIMILARITY: Contains 1 homo B/P domain.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial/
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AF184615; AAF26300.1; -.
DR EMBL; AF184616; AAF26301.1; -.
DR EMBL; AF184617; AAF26302.1; -.
DR HSSP; Q99405; IMPT.
DR MEROPS; S08.OPB; -.
DR InterPro; IPR006212; Furin_repeat.
DR InterPro; IPR009030; Grow_fac_recep.
DR InterPro; IPR000209; Peptidase_S8.
DR InterPro; IPR002884; Peptidase_S8B.
DR InterPro; IPR009020; Protease_1mb1b.
DR Pfam; PF01483; P_piroprotein; 1.
DR Pfam; PF00082; Peptidase_S8; 1.
DR PRINTS; PR00723; SUBTILISIN.
DR ProDom; PD006717; P_domain; 1.
DR SMART; SM00261; FU; 17.
DR PROSITE; PS00136; SUBTILASE_ASP, FALSE_NEG.
DR PROSITE; PS00137; SUBTILASE_HIS; 1.
DR PROSITE; PS00138; SUBTILASE_SER; 1.
KM Hydroxylase: Serine protease; Zymogen; Signal;
KM Cleavage on pair of basic residues; Repeat; Alternative splicing;
KM Transmembrane.
FT SIGNAL 1 25
FT PROPEP 26 110
FT CHAIN 111 1696
FT DOMAIN 111 1618
FT TRANSMEM 1619 1639
FT DOMAIN 1640 1696
FT DOMAIN 111 488
FT DOMAIN 496 637
FT DOMAIN 664 1649
FT SITE 110 111
FT ACT_SITE 192 192
FT ACT_SITE 203 203
FT ACT_SITE 407 407
FT CARBOHYD 246 246
FT CARBOHYD 529 529
FT CARBOHYD 885 885
FT VARSPLIC 1259 1323

```

```

FT FT PREVALEALIGHLYSLDVPDQSNPPDVLGADRRL
FT FT TTATSAAGCA (in isoform C).
FT FT /Ftld=VSP_005442.
FT FT Missing (in isoform C).
FT FT /Ftld=VSP_005443.
FT FT CHPTCKESDSEYDDTCTACNDGFLTLTADSCAGCPGQFL
FT FT HHGDCBCHRECKTC -> IARCDDRDRSCDLVRFNFC
FT FT VRRYFVRCCGCTKLYMEDRPMRGGSSOPTQGRN (in
FT FT isoform A).
FT FT /Ftld=VSP_005444.
FT FT Missing (in isoform A).
FT FT /Ftld=VSP_005445.
SQ SEQUENCE 1696 AA; 188410 MW; 281CBEL784257CDB CRC64;

Query Match 61.8%; Score 39; DB 1; Length 1696;
Best Local Similarity 100.0%; Pred. No. 72;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 DVCQDC 7
Db 775 DVCQDC 780

RESULT 9
CMG2_SCHPO STANDARD; PRT; 355 AA.
AC P32434;
DT 01-OCT-1993 (Rel. 27, Created)
DT 01-OCT-1993 (Rel. 27, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE Type I protein geranylgeranyltransferase beta subunit (BC 2.5.1.59)
DE (Type I protein geranylgeranyltransferase beta subunit) (GGTase-I-
DE beta) (PGGT).
GN CMG2 OR SPAC2B1P5_04C.
OS Schizosaccharomyces pombe (Fission yeast).
OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
OC Schizosaccharomycetales; Schizosaccharomycetaceae;
OC Schizosaccharomycetes.
OX NCBI_TaxID=4896;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=94085400; PubMed=8262067;
RA Diaz M., Sanchez Y., Bennett T., Dun C.R., Godoy C., Tamanoi F.,
RA Duran A., Perez P.;
RT "The Schizosaccharomycetes pombe cmg2+ gene codes for the beta subunit
RT of a geranylgeranyltransferase type I required for beta-glucan
RT synthesis.",
RL EMO J. 12:5245-5254 (1993).
RN [2]
RP SEQUENCE FROM N.A.
RX STRAIN=972;
RX MEDLINE=21848401; PubMed=11859360;
RA Wood V., Gwilliam R., Rajandream M.A., Lyne M., Lyne R., Stewart A.,
RA Sgouras J., Peat N., Hayles J., Baker S., Basham D., Bowman S.,
RA Brooks K., Brown D., Brown S., Chillingworth T., Churcher C.M.,
RA Collins M., Connor R., Cronin A., Davis P., Fellwell T., Fraser A.,
RA Gentles S., Goble A., Hamlin N., Harris D., Hidalgo J., Hodgson G.,
RA Holtroyd S., Hornby T., Howarth S., Huckle E.J., Hunt S., Jagels K.,
RA James K., Jones L., Jones M., Leathers S., McDonald S., McLean J.,
RA Mooney P., Moule S., Mungall K., Murphy L., Niblett D., Odell C.,
RA Oliver K., O'Neill S., Pearson D., Quail M.A., Rabinowitsch E.,
RA Rutherford K., Rutter S., Saunders D., Seeger K., Sharp S.,
RA Skelton J., Simmonds M., Squares R., Squares S., Stevens K.,
RA Taylor K., Taylor R.G., Tivey A., Walsh S.V., Warren T., Whitehead S.,
RA Woodward J., Voiclaert G., Aert R., Robben J., Grymopiez B.,
RA Welljams I., Vansetrel E., Rieger M., Schelefer M., Mueller-Auer S.,
RA Gabel C., Fuchs M., Fritze C., Holzer E., Moestl D., Hilbert H.,
RA Borzym K., Langer I., Beck A., Leirach H., Reinhardt R., Pohl T.M.,
RA Eger P., Zimmermann W., Wedler H., Wambutt R., Purnelle B.,
RA Goffeau A., Cadieu S.J., Dreno S., Gloux S., Lelaure V., Mottier S.,
RA Lucas M., Rochet M., Galliardin C., Tallade V.A., Garzon A., Thode G.,
RA Daga R.R., Cruzado L., Jimenez J., Sanchez M., del Rey F., Benito J.,

```

```

RA Dominguez A., Revuelta J.L., Moreno S., Armstrong J., Forsburg S.L.,
RA Cerutti L., Lowe T., McCombie W.R., Paulsen I., Potashkin J.,
RA Shpakovski G.V., Uesery D., Barrell B.G., Nurse P.;
RT "The genome sequence of Schizosaccharomyces pombe.";
RL Nature 415:871-880 (2002).
CC -! FUNCTION: Catalyzes the transfer of a geranyl-geranyl moiety from
CC geranyl-geranyl pyrophosphate to proteins having the C-terminal C-
CC A-A-L where A is an aliphatic amino acid. In particular it
CC modifies the GTP-binding component of the 1,3-beta-D-glucan
CC synthase.
CC -! CATALYTIC ACTIVITY: Geranylgeranyl diphosphate + protein-cysteine
CC -! -> S-geranylgeranyl-protein + diphosphate.
CC -! COFACTOR: Binds 1 zinc ion per subunit (by similarity).
CC -! SUBUNIT: Heterodimer of an alpha and a beta subunit.
CC -! SIMILARITY: Belongs to the protein prenyltransferase beta subunit
CC family.
CC -! SIMILARITY: Contains 4 PFTB repeats.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; Z12155; CA78143.1; -.
DR EMBL; AL163071; CAB86347.1; -.
DR PIR; S4186; S4186.
DR HSSP; Q02293; 1FT1.
DR GenDB Spombe; SPAC2B1P5_04C; -.
DR InterPro; IPR001330; Prenyltrans.
DR InterPro; IPR008930; Terp_cyc_toroid.
DR Pfam; PR00432; Prenyltrans; 4.
KM Transferase; Prenyltransferase; Repeat; zinc.
FT REPEAT 119 163 PFTB 1.
FT REPEAT 169 210 PFTB 2.
FT REPEAT 234 275 PFTB 3.
FT REPEAT 282 324 PFTB 4.
FT METAL 260 260 ZINC (BY SIMILARITY).
FT METAL 262 262 ZINC (BY SIMILARITY).
FT METAL 312 312 ZINC (BY SIMILARITY).
SQ SEQUENCE 355 AA; 40023 MW; 811A3CBFF067071 CRC64;

Query Match 60.3%; Score 38; DB 1; Length 355;
Best Local Similarity 62.5%; Pred. No. 24;
Matches 5; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 2 DVCQDCIQ 9
Db 343 NICRDCIQ 350

RESULT 10
Y777_TREPA STANDARD; PRT; 74 AA.
AC O83756;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Hypothetical protein TP0777.
GN TP0777.
OS Treponema pallidum.
OC Bacteria; Spirochaetes; Spirochaetales; Spirochaetaceae; Treponema.
OX NCBI_TaxID=160;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=Nichols;
RX MEDLINE=98332770; PubMed=9665876;
RX Fraser C.M., Norris S.J., Weinstein G.M., White O., Sutton G.G.,
RX Dodson R., Gwin M., Hickey E.K., Clayton R., Ketchum K.A.,
RX Sodergren E., Hardham J.M., McLeod M.P., Salzberg S., Peterson J.,
RA Khailak H., Richardson D., Howell J.K., Chidambaram M., Uitterback T.,

```

RA McDonald L., Attiach P., Bowman C., Cotton M.D., Fujii C., Garland S.,  
RA Hatch B., Horst K., Roberts K., Sandusky M., Weidman J., Smith H.O.,  
RA Venter J.C.,  
RT "Complete genome sequence of *Treponema pallidum*, the syphilis  
RT spirochete.";  
RT Science 281:375-388(1998).  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation-  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
DR EMBL, AE001249; AAC65750.1; --  
DR PIR, B71282; B71282.  
DR TIGR, TP0777; --  
KW Hypothetical protein; Complete proteome.  
SQ SEQUENCE 74 AA; 9018 MW; 299EB9E2CEC2D55 CRC64;  
  
Query Match 58.7%; Score 37; DB 1; Length 74;  
Best Local Similarity 50.0%; Pred. No. 7.7;  
Matches 5; Conservative 3; Mismatches 2; Indels 0; Gaps 0;  
  
OY 2 DVCDCDCIOMV 11  
DB 29 DVCDCDCESI 38  
  
RESULT 11  
ID NADA\_HALNT1 STANDARD; PRT; 374 AA.  
AC Q9H2Z1;  
DT 28-FEB-2003 (Rel. 41, Created)  
DT 28-FEB-2003 (Rel. 41, Last sequence update)  
DT 28-FEB-2003 (Rel. 41, Last annotation update)  
DE Quinolinate synthetase A.  
GN NADA OR VNG1882G.  
OS Halobacterium sp. (strain NRC-1 / ATCC 700922 / JCM 11081).  
OC Archaea; Euryarchaeota; Halobacteria; Halobacteriales;  
OC Halobacteriaceae; Halobacterium.  
OX NCBI\_TaxID=64091;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=20504483; PubMed=11016950;  
RA Ng W.V., Kennedy S.P., Mahairas G.G., Bergquist B., Pan M.,  
RA Shukla H.D., Laskey S.R., Baliga N.S., Thorsson V., Shroga J.,  
RA Swartzell S., Weir D., Hall J., Dahl T.A., Welti R., Goo Y.A.,  
RA Leitbauer B., Keller K., Cruz R., Danson M.J., Hough D.W.,  
RA Maddocks D.G., Jablonski P.E., Krebs M.P., Angevine C.M., Dale H.,  
RA Isenbarger T.A., Peck R.F., Pohlischer M., Spudis J.L., Jung K.-H.,  
RA Alan M., Freitas T., Hou S., Daniels C.J., Dennis P.P., Omer A.D.,  
RA Ehardt H., Lowe T.M., Liang P., Riley M., Hood L., Dassarma S.;  
RT "Genome sequence of Halobacterium species NRC-1.";  
RT Proc. Natl. Acad. Sci. U.S.A. 97:12176-12181(2000).  
RL -----  
CC -1- FUNCTION: Catalyzes the condensation of iminoaspartate with  
CC dihydroxyacetone phosphate to form quinolinate (By similarity).  
CC -1- PATHWAY: NAD biosynthesis; aspartate to NADM; second step.  
CC -1- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).  
CC -1- SIMILARITY: Belongs to the quinolinate synthetase A family.  
CC Subfamily 3.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation-  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
DR EMBL, AB005087; AAG20079.1; --  
DR PIR, C84339; C84339.

DR HAMAP; MF 00569; -; 1.  
DR InterPro; IPR003473; NADA.  
DR Pfam; PF02445; NADA; 1.  
DR TIGRFAMs; TIGR00550; nada; 1.  
KW Pyridine nucleotide biosynthesis; Complete proteome.  
SQ SEQUENCE 374 AA; 41305 MW; A52B4985169498B2 CRC64;  
  
Query Match 58.7%; Score 37; DB 1; Length 374;  
Best Local Similarity 60.0%; Pred. No. 37;  
Matches 6; Conservative 0; Mismatches 4; Indels 0; Gaps 0;  
  
OY 1 GDVDCDCIOMV 10  
DB 321 GDVDCDCNAM 330  
  
RESULT 12  
ID U183\_HUMAN STANDARD; PRT; 422 AA.  
AC Q9BSU1;  
DT 28-FEB-2003 (Rel. 41, Created)  
DT 28-FEB-2003 (Rel. 41, Last sequence update)  
DT 10-OCT-2003 (Rel. 42, Last annotation update)  
DE UPF0183 protein.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Lee Y.M., Kim W.;  
RT "Identification of human homolog of rat lin-10.";  
RT Submitted (Aug-1999) to the EMBL/Genbank/DBJ databases.  
RN [2]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Placenta;  
RX MEDLINE=22388257; PubMed=12477932;  
RA Struhsberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,  
RA Klausner R.D., Collins F.S., Wagner L., Shemmen C.M., Schuler G.D.,  
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bat N.K.,  
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,  
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,  
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,  
RA Brownstein M.J., Usdin T.B., Toshitaki S., Carrinci P., Prange C.,  
RA Rana S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullany S.J.,  
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,  
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,  
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,  
RA Fahney J., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A.,  
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,  
RA Bladesley R.W., Touchman J.W., Green E.D., Dickson M.C.,  
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,  
RA Butcherfield Y.S.N., Krzywinski M.I., Skalska U., Smallie D.E.,  
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;  
RT "Generation and initial analysis of more than 15,000 full-length  
RT human and mouse cDNA sequences.";  
RT Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).  
RL -----  
CC -1- SIMILARITY: Belongs to the UPF0183 family.  
CC -1- CAUTION: Lin-10 has been incorrectly assigned for a number of  
CC years, therefore what is stated in Ref.1 as a lin-10 homolog as  
CC nothing to do with the real lin-10.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation-  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
DR EMBL, AF176088; AAK96888.1; --  
DR EMBL, BC004556; AAH04556.1; --  
DR InterPro; IPR005373; UPF0183.

DR Pfam; PF03676; UPF0183; 1.  
SQ SEQUENCE 422 AA; 47524 MW; CEA42442E45A7E35 CRC64;

Query Match 58.7%; Score 37; DB 1; Length 422;  
Best Local Similarity 54.5%; Pred. No. 41;  
Matches 6; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 GDVQCDCIOMV 11  
DB 242 GDSCQDVLMSL 252

RESULT 13

ID U183\_MOUSE STANDARD; PRT; 422 AA.  
AC Q922R1;  
DT 28-FEB-2003 (Rel. 41, Created)  
DT 28-FEB-2003 (Rel. 41, Last sequence update)  
DT 10-OCT-2003 (Rel. 42, Last annotation update)  
DE UPF0183 protein.  
OS Mus musculus (Mouse).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
OX NCBI\_TaxId=10990;  
RN [1]

SEQUENCE FROM N.A.  
MEDLINE=22388257; PubMed=12477932;  
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,  
RA Klausner R.D., Collins F.S., Wagner L., Shennan C.M., Schuler G.D.,  
RA Altschul S.F., Zeeberg B., Buetow K.H., Scheffer C.F., Bhat N.K.,  
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,  
RA Datchenko L., Matovina K., Farmer A.A., Rubin G.M., Hong L.,  
RA Stapleton M., Soares M.B., Bonaldo M.P., Casavant T.L., Schaeetz T.E.,  
RA Brownstein M.J., Ueda T.B., Toshiyuki S., Carninci P., Prange C.,  
RA Raha S.S., Loqueanno N.A., Peters G.J., Abramson R.D., Mullany S.J.,  
RA Bosak S.A., McSwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,  
RA Richard S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hultyk S.W.,  
RA Villalón D.K., Murzy D.M., Sodergren E.J., Lu X., Gibbs R.A.,  
RA Fahy J., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A.,  
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,  
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,  
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,  
RA Buterfield Y.S.N., Krzywinski M.I., Skalska U., Smalish D.E.,  
RA Schercher A., Schein J.E., Jones S.J.M., Marra M.A.,  
RT "Generation and initial analysis of more than 15,000 full-length  
human and mouse cDNA sequences.";  
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).  
CC -1 SIMILARITY: Belongs to the UPF0183 family.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

DR EMBL; BC006874; AAH06874.1;  
DR MGD; MGI:2384564; BC006874.  
DR InterPro; IPR005373; UPF0183.  
DR Pfam; PF03676; UPF0183; 1.  
SQ SEQUENCE 422 AA; 47402 MW; 5639E30A26D96F14 CRC64;

Query Match 58.7%; Score 37; DB 1; Length 422;  
Best Local Similarity 54.5%; Pred. No. 41;  
Matches 6; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 GDVQCDCIOMV 11  
DB 242 GDSCQDVLMSL 252

RESULT 14

U183\_RAT STANDARD; PRT; 422 AA.  
ID U183\_RAT  
AC Q08654;  
DT 28-FEB-2003 (Rel. 41, Created)  
DT 28-FEB-2003 (Rel. 41, Last sequence update)  
DT 28-FEB-2003 (Rel. 41, Last annotation update)  
DE UPF0183 protein.  
OS Rattus norvegicus (Rat).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.  
OX NCBI\_TaxId=10116;  
RN [1]

SEQUENCE FROM N.A.  
RA STRAIN=Sprague-Dawley;  
RA MEDLINE=96153816; PubMed=9480860;  
RA Ide N., Hirao K., Hata Y., Takeuchi M., Irie M., Yao I., Deguchi M.,  
RA Toyoda A., Nishiohara H., Mizoguchi A., Takai Y.,  
RT "Molecular cloning and characterization of rat lin-10.";  
RL Biochem. Biophys. Res. Commun. 243:634-638(1998).  
CC -1 SIMILARITY: Belongs to the UPF0183 family.  
CC -1 CAUTION: Lin-10 has been incorrectly assigned for a number of  
CC years, therefore what is stated in Ref.1 as a lin-10 homolog as  
CC nothing to do with the real lin-10.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

DR EMBL; U92010; AAB51383.1;  
DR PIR; J02039; J02039.  
DR InterPro; IPR005373; UPF0183.  
DR Pfam; PF03676; UPF0183; 1.  
SQ SEQUENCE 422 AA; 47463 MW; D5B185189FE2E320 CRC64;

Query Match 58.7%; Score 37; DB 1; Length 422;  
Best Local Similarity 54.5%; Pred. No. 41;  
Matches 6; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 GDVQCDCIOMV 11  
DB 242 GDSCQDVLMSL 252

RESULT 15

ID CLPE\_LACLA STANDARD; PRT; 748 AA.  
AC Q9C109;  
DT 16-OCT-2001 (Rel. 40, Created)  
DT 16-OCT-2001 (Rel. 40, Last sequence update)  
DT 28-FEB-2003 (Rel. 41, Last annotation update)  
DE ATP-dependent Clp protease ATP-binding subunit clpE.  
CC CLPE OR L10557.

OS Lactococcus lactis (subsp. lactis) (Streptococcus lactis).  
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae; Lactococcus.  
OX NCBI\_TaxId=1360;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=IL1403;  
RA MEDLINE=11403;  
RA Bologin A., Winkler P., Mauger S., Jallion O., Malarme K.,  
RA Weisenbach J., Ehrlich S.D., Sorokin A.,  
RT "The complete genome sequence of the lactic acid bacterium Lactococcus  
RT lactis sep. lactis 111403.";  
RL Genome Res. 11:731-753(2001).  
CC -1 FUNCTION: COULD BE NECESSARY FOR DEGRADING PROTEINS GENERATED BY  
CC CERTAIN TYPES OF STRESS (By similarity).  
CC -1 INDUCTION: By heat shock (Probable).  
CC -1 SIMILARITY: BELONGS TO THE CLPA/CLPB FAMILY. CLPE SUBFAMILY.

CC -1 SIMILARITY: Contains 1 UVR domain.

```
-----
CC      This SWISS-PROT entry is copyright. It is produced through a collaboration
CC      between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC      the European Bioinformatics Institute. There are no restrictions on its
CC      use by non-profit institutions as long as its content is in no way
CC      modified and this statement is not removed. Usage by and for commercial
CC      entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC      or send an email to license@isb-sib.ch).
CC      -----
DR      EMBL; AE006289; AK04655.1; -.
DR      PIR; E86694; E86694.
DR      InterPro; IPR003593; AAA_ATPase.
DR      InterPro; IPR003599; AAA_ATPase_cent.
DR      InterPro; IPR001270; Chaperin_c1pa/B.
DR      InterPro; IPR001943; UvrB/C.
DR      Pfam; PF00004; AAA; 2.
DR      Pfam; PF02151; Uvr; 1.
DR      PRINTS; PR00300; CLPPROTEASEA.
DR      SMART; SM00382; AAA; 2.
DR      PROSITE; PS00870; CLPAB_1; 1.
DR      PROSITE; PS00871; CLPAB_2; 1.
DR      PROSITE; PS50151; Uvr; 1.
DR      Chaperone; ATP-binding; Heat shock; Zinc-finger; Repeat;
KW      Complete proteome.
KW      DOMAIN 107 354 I.
FT      DOMAIN 359 394 Uvr.
FT      DOMAIN 407 598 II.
FT      ZN_FING 3 32 C4-TYPE (POTENTIAL).
FT      NP_BIND 152 159 ATP (POTENTIAL).
FT      NP_BIND 481 488 ATP (POTENTIAL).
SQ      SEQUENCE 748 AA; 83144 MW; 738ECF1B994B31B7 CRC64;

Query Match      58.7%; Score 37; DB 1; Length 748;
Best Local Similarity 50.0%; Pred. No. 71;
Matches 5; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

QY      2 DVCQDCIQMV 11
DB      27 DLQNCYQIM 36
```

Search completed: May 5, 2004, 13:31:44  
Job time : 1.27161 secs

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 5, 2004, 13:26:23 ; Search time 0.438746 Seconds  
(without alignments)  
2411.659 Million cell updates/sec

Title: US-09-743-684A-19  
Perfect score: 63  
Sequence: 1 GDVQCDCIQMV 11

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 283366 seqs, 96191526 residues

Total number of hits satisfying chosen parameters: 283366

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : PIR\_78.\*

1: pir1.\*

2: pir2.\*

3: pir3.\*

4: pir4.\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	63	100.0	79	2 A49475	cerebroside sulfat
2	63	100.0	527	1 SAHUP	saposin precursor
3	47	74.6	554	1 A28716	saposin precursor
4	47	74.6	557	1 JH0604	saposin precursor
5	40	63.5	46	2 E69830	hypothetical prote
6	40	63.5	1356	2 F84486	probable retroelem
7	39	61.9	320	2 B49810	gene 7 protein - m
8	39	61.9	512	1 VGI07H	envelope glycoprot
9	39	61.9	666	2 T10567	probable serine/th
10	38	60.3	221	2 T46918	hypothetical prote
11	38	60.3	307	2 T40043	probable transport
12	38	60.3	320	2 E70355	conserved hypothet
13	38	60.3	348	2 A57234	lin-44 protein pre
14	38	60.3	355	2 S41686	geranylgeranyltran
15	38	60.3	456	2 T24442	hypothetical prote
16	38	60.3	1672	2 T46237	hypothetical prote
17	38	60.3	2609	2 T40399	probable transport
18	37	58.7	74	2 B71282	hypothetical prote
19	37	58.7	374	2 C48339	quinolinolate synthe
20	37	58.7	413	2 F86847	hypothetical prote
21	37	58.7	422	2 J80239	lin-10 protein - r
22	37	58.7	429	2 T41560	phenylalanyl-trna
23	37	58.7	748	2 E86694	ATP-dependent prot
24	37	58.7	5147	1 IJFFTM	cadherin-related t
25	36	57.1	47	2 I48943	cellular disintegr
26	36	57.1	111	2 B71239	hypothetical prote
27	36	57.1	151	2 S60259	meltrin gamma - mo
28	36	57.1	213	2 F70476	conserved hypothet
29	36	57.1	218	2 AH1388	late competence pr

30	36	57.1	291	2 T20666	hypothetical prote
31	36	57.1	294	2 T20664	hypothetical prote
32	36	57.1	330	2 A29915	teratocarcinoma gl
33	36	57.1	367	2 T24298	hypothetical prote
34	36	57.1	493	2 T29833	hypothetical prote
35	36	57.1	511	2 S10527	endoglucanase B pr
36	36	57.1	655	2 JC7850	disintegrin and me
37	36	57.1	660	2 S71949	metalloproteinase
38	36	57.1	778	2 D85055	probable polypro
39	36	57.1	1363	2 T47492	copia-like polypro
40	35.5	56.3	1416	2 E88550	protein ZC84.1 (im
41	35.5	56.3	2844	2 S28291	hypothetical prote
42	35	55.6	101	2 S62337	L7i-8 protein - fr
43	35	55.6	103	2 A47337	rhodniin - Rhodniu
44	35	55.6	108	2 A86675	hypothetical prote
45	35	55.6	127	2 F87303	cheU protein (impo

ALIGNMENTS

RESULT 1

A49475

cerebroside sulfat activator - pig

C:Species: Sus scrofa domestica (domestic pig)

C:Date: 24-Feb-1994 #sequence\_revision 18-Nov-1994 #text\_change 17-May-1996

C:Accession: A49475

R:Stevens, R.L.; Faull, K.F.; Conklin, K.A.; Green, B.N.; Fluharty, A.L.

Biochemistry 32, 4051-4059, 1993

A:Title: Porcine cerebroside sulfate activator: further structural characterization and

A:Reference number: A49475; MUID:93229506; PMID:8471613

A:Accession: A49475

A:Status: preliminary

A:Molecule type: protein

A:Residues: 1-79 <STR>

A:Experimental source: kidney

A:Note: sequence extracted from NCBI backbone (NCBIP:129597)

C:Superfamily: saposin; saposin repeat homology

F:1-79/domain: saposin repeat homology <SAP>

Query Match 100.0%; Score 63; DB 2; Length 79;

Best Local Similarity 100.0%; Pred. No. 0.0019;

Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GDVQCDCIQMV 11

Db 1 GDVQCDCIQMV 11

RESULT 2

SAHUP

saposin precursor [validated] - human

N:Alternate names: cerebroside sulfate activator protein; co-beta-glucosidase; component

ein (SAP); sphingolipid activator protein A2; sulfate sulfatase activator protein

N:Contains: prosaposin; saposin A; saposin B; saposin C; saposin D

C:Species: Homo sapiens (man)

C:Date: 30-Jun-1992 #sequence\_revision 17-Nov-1995 #text\_change 08-Dec-2000

C:Accession: JX0061; A57368; A42003; B42003; C42003; D42003; A30367; S34740; S36140; S36

0226; I37265; I37264

R:Nakano, T.; Sandhoff, K.; Stuenkel, J.; Christomanou, H.; Suzuki, K.

J. Biochem. 105, 152-154, 1989

A:Title: Structure of full-length cDNA coding for sulfate activator, a Co-beta-glucosi

A:Reference number: JX0061; MUID:89255151; PMID:2498298

A:Accession: JX0061

A:Molecule type: mRNA

A:Residues: 1-527 <NAK>

A:Cross-references: GB:P00422; NID:g2220063; PIDN:BAA00321.1; PID:g2220064

A:Note: alternative splice form 1

A:Accession: A57368

A:Molecule type: mRNA

A:Residues: 1-259,263-527 <NA2>

A:Cross-references: GB:J03015; GB:J03086; NID:g337755; PIDN:AAB59494.1; PID:g337756

A:Note: alternative splice form 2

R;Rorman, E.G.; Scheinlker, V.; Grabowski, G.A.  
 Genomics 13, 312-318, 1992  
 A;Title: Structure and evolution of the human prosaposin chromosomal gene.  
 A;Reference number: A42003; MUID:92307663; PMID:1612590  
 A;Accession: A42003  
 A;Molecule type: DNA  
 A;Residues: 50-140 <ROR>  
 A;Cross-references: GB:M86181  
 A;Note: sequence extracted from NCBI backbone (NCBIN:107235, NCBIP:107236)  
 A;Accession: B42003  
 A;Molecule type: DNA  
 A;Residues: 185-259;263-276 <RO2>  
 A;Note: sequence extracted from NCBI backbone (NCBIN:107235, NCBIP:107237)  
 A;Accession: C42003  
 A;Molecule type: DNA  
 A;Residues: 305-393 <RO3>  
 A;Note: sequence extracted from NCBI backbone (NCBIN:107235, NCBIP:107238); sequence inc  
 A;Accession: D42003  
 A;Molecule type: DNA  
 A;Residues: 399-487 <RO4>  
 A;Note: sequence extracted from NCBI backbone (NCBIN:107235, NCBIP:107239); sequence inc  
 R;Rorman, E.G.; Grabowski, G.A.  
 Genomics 5, 486-492, 1989  
 A;Title: Molecular cloning of a human co-beta-glucosidase cDNA: evidence that four sphin  
 A;Reference number: A30367; MUID:90129043; PMID:2515150  
 A;Accession: A30367  
 A;Molecule type: mRNA  
 A;Residues: 1-259;263-527 <RO5>  
 A;Cross-references: GB:J03077; NID:G1832230; PIDN:AAA52560.1; PID:G1832231  
 A;Note: alternative splice form 2  
 R;Hiraiwa, M.; O'Brien, J.S.; Kishimoto, Y.; Galdzicka, M.; Fluharty, A.L.; Gimms, E.I.;  
 Arch. Biochem. Biophys. 304, 110-116, 1993  
 A;Title: Isolation, characterization, and proteolysis of human prosaposin, the precursor  
 A;Reference number: S34740; MUID:93311991; PMID:8323276  
 A;Accession: S34740  
 A;Molecule type: protein  
 A;Residues: 17-24;165-172;180-189;301-305 <HTR>  
 R;Tyynelae, J.; Palmer, D.N.; Baumann, M.; Haltia, M.  
 FEBS Lett. 330, 8-12, 1993  
 A;Title: Storage of saposins A and D in infantile neuronal ceroid-lipofuscinosis.  
 A;Reference number: S36140; MUID:93380576; PMID:8370464  
 A;Accession: S36140  
 A;Molecule type: protein  
 A;Residues: 'XX', 62, 'X', 64-65, 'X', 67-79, 'X', 81-84 <TYV>  
 A;Note: saposin A  
 A;Accession: S36141  
 A;Molecule type: protein  
 A;Residues: 'XXX', 413-414, 'X', 416-428, 'X', 430-434 <TY2>  
 R;Holtzschmidt, H.; Sandhoff, K.; Kwon, H.Y.; Harzer, K.; Nakano, T.; Suzuki, K.  
 J. Biol. Chem. 266, 7556-7560, 1991  
 A;Title: Sulfatide activator protein. Alternative splicing that generates three mRNAs an  
 A;Reference number: S36988; MUID:91210267; PMID:2019586  
 A;Accession: S36988  
 A;Status: nucleic acid sequence not shown; translation not shown  
 A;Molecule type: mRNA  
 A;Residues: 1-240, 'S', 242-527 <HO1>  
 A;Cross-references: EMBL:M60255; NID:G337759; PIDN:AAA36594.1; PID:G337760  
 A;Note: the nucleotide sequence was submitted to the EMBL Data Library, June 1991  
 A;Accession: S36989  
 A;Status: nucleic acid sequence not shown; translation not shown  
 A;Molecule type: mRNA  
 A;Residues: 1-240, 'S', 242-259, 263-527 <HO2>  
 A;Cross-references: EMBL:M60257; NID:G337764; PIDN:AAA36595.1; PID:G337765  
 A;Note: the nucleotide sequence was submitted to the EMBL Data Library, June 1991  
 A;Accession: S36990  
 A;Status: nucleic acid sequence not shown; translation not shown  
 A;Molecule type: mRNA  
 A;Residues: 1-240, 'S', 242-259, 261-527 <HO3>  
 A;Cross-references: EMBL:M60258; NID:G337766; PIDN:AAA36596.1; PID:G337767  
 A;Note: the nucleotide sequence was submitted to the EMBL Data Library, June 1991

A;Note: cerebroside sulfate activator protein mutant MU-6; corresponds to alternative sp  
 R;Kondo, K.; Hineno, T.; Sano, A.; Kakimoto, Y.  
 Biochem. Biophys. Res. Commun. 181, 286-292, 1991  
 A;Title: Isolation and characterization of prosaposin from human milk.  
 A;Reference number: PS0330; MUID:92068206; PMID:1958198  
 A;Accession: PS0330  
 A;Molecule type: protein  
 A;Residues: 17-24, 'X', 26 <KON>  
 A;Experimental source: milk  
 R;Kretz, K.A.; Carson, G.S.; Morimoto, S.; Kishimoto, Y.; Fluharty, A.L.; O'Brien, J.S.  
 Proc. Natl. Acad. Sci. U.S.A. 87, 2541-2544, 1990  
 A;Title: Characterization of a mutation in a family with saposin B deficiency: a glycosyl  
 A;Reference number: A35985; MUID:90207231; PMID:2320574  
 A;Accession: A35985  
 A;Molecule type: mRNA  
 A;Residues: 213-221 <KRE>  
 A;Cross-references: GB:M32221  
 A;Accession: B35985  
 A;Status: nucleic acid sequence not shown; translation not shown  
 A;Molecule type: mRNA  
 A;Residues: 1-259;263-527 <KR2>  
 A;Cross-references: GB:M32221; NID:G337761; PIDN:AAA60303.1; PID:G337762  
 A;Experimental source: lymphoblast  
 A;Accession: C35985  
 A;Molecule type: mRNA  
 A;Residues: 213-216, 'I', 218-221 <KR3>  
 A;Note: sequence from patients with activator-deficient metachromatic leukodystrophy; th  
 R;Fuerst, W.; Schubert, J.; Machleidt, W.; Meyer, H.E.; Sandhoff, K.  
 Eur. J. Biochem. 192, 709-714, 1990  
 A;Title: The complete amino-acid sequences of human ganglioside GM2 activator protein an  
 A;Reference number: S13195; MUID:91006165; PMID:2209618  
 A;Accession: S13196  
 A;Molecule type: protein  
 A;Residues: 195-259;263-277 <FUE>  
 R;Morimoto, S.; Martin, B.M.; Yamamoto, Y.; Kretz, K.A.; O'Brien, J.S.; Kishimoto, Y.  
 Proc. Natl. Acad. Sci. U.S.A. 86, 3389-3393, 1989  
 A;Title: Saposin A: second cerebroside activator protein.  
 A;Reference number: A32784; MUID:89240739; PMID:2717620  
 A;Accession: A32784  
 A;Molecule type: protein  
 A;Residues: 60-84;86-107;109-119;125-134 <MOR>  
 R;O'Brien, J.S.; Kretz, K.A.; Dewji, N.; Wenger, D.A.; Esch, F.; Fluharty, A.L.  
 Science 241, 1098-1101, 1988  
 A;Title: Coding of two sphingolipid activator proteins (SAP-1 and SAP-2) by same genetic  
 A;Reference number: A41240; MUID:88321660; PMID:2842863  
 A;Accession: A41240  
 A;Molecule type: mRNA  
 A;Residues: 'GSSR', 18-259, 263-299, 'D', 301-302, 'D', 304-527 <OAB>  
 A;Cross-references: GB:J03086  
 R;Dewji, N.N.; Wenger, D.A.; O'Brien, J.S.  
 Proc. Natl. Acad. Sci. U.S.A. 84, 8652-8656, 1987  
 A;Title: Nucleotide sequence of cloned cDNA for human sphingolipid activator protein 1 p  
 A;Reference number: S02289; MUID:88068647; PMID:2825202  
 A;Accession: S02289  
 A;Status: significant sequence differences  
 A;Molecule type: mRNA  
 A;Cross-references: EMBL:J03015  
 A;Note: this sequence corrected by A41240  
 A;Note: part of this sequence, including the amino end of the mature protein, was determ  
 R;Klein-schmidt, T.; Christomanou, H.; Braunitzer, G.  
 Biol. Chem. Hoppe-Seyler 369, 1361-1365, 1988  
 A;Title: Complete amino-acid sequence of the naturally occurring A(2) activator protein f  
 A;Reference number: S02028; MUID:89207118; PMID:3242555  
 A;Accession: S02028  
 A;Molecule type: protein  
 A;Residues: 195-259;263-276 <KLE>  
 R;Fuerst, W.; Machleidt, W.; Sandhoff, K.  
 Biol. Chem. Hoppe-Seyler 369, 317-328, 1988  
 A;Title: The precursor of sulfate activator protein is processed to three different pr  
 A;Reference number: S00813; MUID:89000190; PMID:3048308  
 A;Accession: S00813  
 A;Molecule type: protein  
 A;Residues: 410-487 <FU2>



R:Kleinschmidt, T.; Christomanou, H.; Braunitzer, G.  
Biol. Chem. Hoppe-Seyler 368, 1571-1578, 1987  
A:Title: Complete amino-acid sequence and carbohydrate content of the naturally occurring  
A:Reference number: S00226; MUID:88163077; PMID:3442600  
A:Accession: S00226  
A:Molecule type: protein  
A:Residues: 314-393 <KL2>  
R:Vaccaro, A.M.; Salvio, R.; Barca, A.; Tatti, M.; Ciaffoni, F.; Maras, B.; Siciliano, J.  
J. Biol. Chem. 270, 9953-9960, 1995  
A:Title: Structural analysis of saposin C and B. Complete localization of disulfide bridge  
A:Reference number: A57297; MUID:95247790; PMID:7730378  
A:Contents: annotation; disulfide bonds; glycosylation  
R:Holtzman, H.; Sandhoff, K.; Fuerst, W.; Kwon, H.Y.; Schnabel, D.; Suzuki, K.  
FEBS Lett. 280, 267-270, 1991  
A:Title: The organization of the gene for the human cerebroside sulfate activator protein  
A:Reference number: I37264; MUID:91192146; PMID:2013321  
A:Accession: I37265  
A:Status: translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 59-125 <RES>  
A:Cross-references: EMBL:X57107; NID:G30234; PIDN:CAA40391.1; PID:G30235  
A:Accession: I37264  
A:Status: translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 307-516 <RE2>  
A:Cross-references: EMBL:X57108; NID:G30232; PIDN:CAA40392.1; PID:G1565257  
A:Note: sequence revised relative to PID:G30233 (corrected coding region)  
C:Genetics:  
A:Gene: GDB:PSAP; GLRA  
A:Cross-references: GDB:120366; OMIM:176801  
A:Map position: 10q22.1-10q22.1  
A:Introns: 83/3; 338/3; 401/1; 453/3; 480/3  
A:Note: defects in this gene may cause variant Gaucher disease, variant Tay-Sachs disease  
A:Note: list of introns is incomplete

Query Match 100.0%; Score 63; DB 1; Length 527;  
Best Local Similarity 100.0%; Pred. No. 0.0091;  
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GDVQCDCIQMV 11  
|||||  
Db 195 GDVQCDCIQMV 205  
|||||

RESULT 3  
A28716  
saposin precursor - rat  
N:Alternate names: cerebroside sulfate activator protein; co-beta-glucosidase; component  
ein (SAP); sphingolipid activator protein A2; sulfated glycoprotein 1; sulfatide sulfate  
N:Contains: prosaposin; saposin A; saposin B; saposin C; saposin D  
C:Species: Rattus norvegicus (Norway rat)  
C:Date: 10-Sep-1999 #sequence\_revision 10-Sep-1999 #text\_change 10-Sep-1999  
C:Accession: A28716  
R:Collard, M.W.; Sylvester, S.R.; Tsuruta, J.K.; Griswold, M.D.  
Biochemistry 27, 4557-4564, 1988  
A:Title: Biosynthesis and molecular cloning of sulfated glycoprotein 1 secreted by rat  
A:Reference number: A28716; MUID:89000647; PMID:3048385  
A:Accession: A28716  
A:Molecule type: mRNA  
A:Residues: 1-554 <COL>  
A:Cross-references: GB:M19936; NID:G206904; PIDN:AAA42136.1; PID:G206905  
A:Note: parts of this sequence, including the amino end of the mature protein, were deter  
C:Function:  
A:Description: saposins bind sphingolipids, form hydrophilic complexes and make them acc  
A:Pathway: sphingolipid catabolism  
A:Note: saposins A and C (SAP-2) activate hydrolysis of glucocerebroside by beta-glucosy  
A:Note: saposin B (SAP-1) activates hydrolysis of galactocerebroside sulfate by arylsul  
A:Note: saposin D activates hydrolysis of sphingomyelin by sphingomyelin phosphodiester  
C:Superfamily: saposin; saposin repeat homology  
C:Keywords: alternative splicing; glycoprotein; lysosomal storage disease; lysosome; sph  
F:1-16/Domain: signal sequence #status predicted <SIG>  
F:17-554/Product: prosaposin #status predicted <PRO>

Query Match 100.0%; Score 63; DB 1; Length 527;  
Best Local Similarity 100.0%; Pred. No. 0.0091;  
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GDVQCDCIQMV 11  
|||||  
Db 195 GDVQCDCIQMV 205  
|||||

RESULT 3  
A28716  
saposin precursor - rat  
N:Alternate names: cerebroside sulfate activator protein; co-beta-glucosidase; component  
ein (SAP); sphingolipid activator protein A2; sulfated glycoprotein 1; sulfatide sulfate  
N:Contains: prosaposin; saposin A; saposin B; saposin C; saposin D  
C:Species: Rattus norvegicus (Norway rat)  
C:Date: 10-Sep-1999 #sequence\_revision 10-Sep-1999 #text\_change 10-Sep-1999  
C:Accession: A28716  
R:Collard, M.W.; Sylvester, S.R.; Tsuruta, J.K.; Griswold, M.D.  
Biochemistry 27, 4557-4564, 1988  
A:Title: Biosynthesis and molecular cloning of sulfated glycoprotein 1 secreted by rat  
A:Reference number: A28716; MUID:89000647; PMID:3048385  
A:Accession: A28716  
A:Molecule type: mRNA  
A:Residues: 1-554 <COL>  
A:Cross-references: GB:M19936; NID:G206904; PIDN:AAA42136.1; PID:G206905  
A:Note: parts of this sequence, including the amino end of the mature protein, were deter  
C:Function:  
A:Description: saposins bind sphingolipids, form hydrophilic complexes and make them acc  
A:Pathway: sphingolipid catabolism  
A:Note: saposins A and C (SAP-2) activate hydrolysis of glucocerebroside by beta-glucosy  
A:Note: saposin B (SAP-1) activates hydrolysis of galactocerebroside sulfate by arylsul  
A:Note: saposin D activates hydrolysis of sphingomyelin by sphingomyelin phosphodiester  
C:Superfamily: saposin; saposin repeat homology  
C:Keywords: alternative splicing; glycoprotein; lysosomal storage disease; lysosome; sph  
F:1-16/Domain: signal sequence #status predicted <SIG>  
F:17-554/Product: prosaposin #status predicted <PRO>

F:60-143/Product: saposin A #status predicted <SAPA>  
F:189-280/Domain: saposin repeat homology <SAP2>  
F:194-273/Product: saposin B #status predicted <SAB1>  
F:306-397/Domain: saposin repeat homology <SAP3>  
F:310-389/Product: saposin C #status predicted <SAPC>  
F:431-522/Domain: saposin repeat homology <SAP4>  
F:437-514/Product: saposin D #status predicted <SAPD>  
F:63-138,66-132,94-106,439-512,442-506,470-481/Disulfide bonds: #status predicted  
F:80,214,331,456/Binding site: carbohydrate (Asn) (covalent) #status predicted  
F:197-270,200-264,229-240,314-387,317-381,345-356/Disulfide bonds: #status predicted

Query Match 74.6%; Score 47; DB 1; Length 554;  
Best Local Similarity 70.0%; Pred. No. 3.6;  
Matches 7; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

Qy 2 DVCQDCIQMV 11  
|||||  
Db 195 DVCQDCMKLV 204  
|||||

RESULT 4  
JH0604  
saposin precursor - mouse  
N:Alternate names: cerebroside sulfate activator protein; co-beta-glucosidase; component  
ein (SAP); sphingolipid activator protein A2; sulfated glycoprotein 1; sulfatide sulfate  
N:Contains: prosaposin; saposin A; saposin B; saposin C; saposin D  
C:Species: Mus musculus (house mouse)  
C:Date: 10-Sep-1999 #sequence\_revision 10-Sep-1999 #text\_change 10-Sep-1999  
C:Accession: JH0604  
R:Tsuada, M.; Sakiyama, T.; Endo, H.; Kitagawa, T.  
Biochem. Biophys. Res. Commun. 184, 1266-1272, 1992  
A:Title: The primary structure of mouse saposin.  
A:Reference number: JH0604; MUID:92272718; PMID:1590788  
A:Accession: JH0604  
A:Molecule type: mRNA  
A:Residues: 1-557 <TSU>  
A:Cross-references: GB:S36200; NID:G249386; PIDN:AA22175.1; PID:G249387  
A:Experimental source: liver  
C:Function:  
A:Description: saposins bind sphingolipids, form hydrophilic complexes and make them acc  
A:Pathway: sphingolipid catabolism  
A:Note: saposins A and C (SAP-2) activate hydrolysis of glucocerebroside by beta-glucosy  
A:Note: saposin B (SAP-1) activates hydrolysis of galactocerebroside sulfate by arylsul  
A:Note: saposin D activates hydrolysis of sphingomyelin by sphingomyelin phosphodiester  
C:Superfamily: saposin; saposin repeat homology  
C:Keywords: alternative splicing; glycoprotein; lysosomal storage disease; lysosome; sph  
F:1-16/Domain: signal sequence #status predicted <SIG>  
F:17-557/Product: prosaposin #status predicted <PRO>  
F:55-148/Domain: saposin repeat homology <SAP1>  
F:60-143/Product: saposin A #status predicted <SAPA>  
F:189-283/Domain: saposin repeat homology <SAP2>  
F:194-276/Product: saposin B #status predicted <SAB1>  
F:309-400/Domain: saposin repeat homology <SAP3>  
F:313-392/Product: saposin C #status predicted <SAPC>  
F:434-525/Domain: saposin repeat homology <SAP4>  
F:440-517/Product: saposin D #status predicted <SAPD>  
F:63-138,66-132,94-106,197-273,200-267,229-240,317-390,320-384,348-359,442-515,445-509,4  
F:80,214,334,379,459/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 74.6%; Score 47; DB 1; Length 557;  
Best Local Similarity 70.0%; Pred. No. 3.6;  
Matches 7; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

Qy 2 DVCQDCIQMV 11  
|||||  
Db 195 DVCQDCMKLV 204  
|||||

RESULT 5  
E69830  
hypothetical protein yfH - Bacillus subtilis  
C:Species: Bacillus subtilis  
C:Date: 05-Dec-1997 #sequence\_revision 05-Dec-1997 #text\_change 15-Oct-1999

C;Accession: E69830  
R;Kunst, F.; Ogasawara, N.; Moszer, I.; Albertini, A.M.; Alloni, G.; Azevedo, V.; Bertea  
C.; Bron, S.; Brouillet, S.; Brusch, C.V.; Caldwell, B.; Capuano, V.; Carter, N.M.; Chd  
A.; Ehrlich, S.D.; Emerson, P.T.; Entian, K.D.; Errington, J.; Fabret, C.; Ferrari, E.  
Nature 390, 249-256, 1997  
A;Authors: Foulger, D.; Frit, C.; Fujita, M.; Fujita, Y.; Fuma, S.; Galizzi, A.; Galle  
iech, J.; Harwood, C.R.; Henaut, A.; Hilbert, H.; Holsappel, S.; Hosono, S.; Hullo, M.F.  
Koeter, P.; Konigstein, G.; Krogh, S.; Kumano, M.; Kurita, K.; Lapidus, A.; Lardinois,  
A;Authors: Lauber, J.; Lazarevic, V.; Lee, S.M.; Levine, A.; Liu, H.; Masuda, S.; Maueel  
Y, M.; Ogawa, K.; Ogiwara, A.; Oudega, B.; Park, S.H.; Parro, V.; Pohl, T.M.; Portetelle  
Rieger, M.; Rivolta, C.; Rocha, E.; Roche, B.; Rose, M.; Sadaie, Y.; Sato, T.; Scanlon,  
A;Authors: Schleich, S.; Schroeter, R.; Scoffone, P.; Sekiguchi, J.; Sekowska, A.; Seron  
akeuchi, M.; Tamakoshi, A.; Tanaka, T.; Terpstra, P.; Tognoni, A.; Tosato, V.; Uchiyama,  
T.; Winters, P.; Wipat, A.; Yamamoto, H.; Yamane, K.; Yasumoto, K.; Yata, K.; Yoshida, K  
A;Authors: Yoshikawa, H.F.; Zumstein, E.; Yoshikawa, H.; Danchin, A.  
A;Title: The complete genome sequence of the Gram-positive bacterium Bacillus subtilis.  
A;Reference number: A69580; MUID:98044033; PMID:9384377  
A;Accession: E69830  
A;Status: preliminary; nucleic acid sequence not shown; translation not shown  
A;Molecule type: DNA  
A;Residues: 1-46 <KUN>  
A;Cross-references: GB:Z99109; GB:AL009126; NID:G2633260; PIDN:CAB12863.1; PID:e1183025;  
A;Experimental source: strain 168  
C;Genetics:  
A;Gene: ynfH

Query Match 63.5%; Score 40; DB 2; Length 46;  
Best Local Similarity 50.0%; Pred. No. 6.1;  
Matches 5; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GDVQDCIQM 10  
|:|:|:|:|:  
Db 34 GNICNDCKV 43

RESULT 6  
F84486  
Probable retroelement pol polyprotein [imported] - Arabidopsis thaliana  
C;Species: Arabidopsis thaliana (mouse-ear cress)  
C;Date: 02-Feb-2001 #sequence\_revision 02-Feb-2001 #text\_change 16-Feb-2001  
C;Accession: F84486  
R;Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.;  
M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; VanAken, S.E.; Unayam, L.; Tallon, L.  
euss, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter, J  
Nature 402, 761-768, 1999  
A;Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.  
A;Reference number: A84420; MUID:20083487; PMID:10617197  
A;Accession: F84486  
A;Status: preliminary  
A;Molecule type: DNA  
A;Residues: 1-1356 <STO>  
A;Cross-references: GB:AE002093; NID:G4895171; PIDN:AAD32759.1; GSPDB:GN00139  
C;Genetics:  
A;Gene: Atg07550  
A;Map position: 2  
C;Superfamily: retrovirus-related polyprotein

Query Match 63.5%; Score 40; DB 2; Length 1356;  
Best Local Similarity 85.7%; Pred. No. 1e+02;  
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 2 DVQDCI 8  
|:|:|:|:  
Db 471 DVCEDCI 477

RESULT 7  
B49810  
gene 7 protein - murine rotavirus B  
C;Species: group B rotavirus  
A;Note: strain IDIR (infectious diarrhea of infant rats)  
C;Date: 01-Dec-1995 #sequence\_revision 01-Dec-1995 #text\_change 18-Jul-2001  
C;Accession: B49810

R;Eiden, J.J.  
Virology 199, 212-218, 1994  
A;Title: Expression and sequence analysis of gene 7 of the IDIR agent (group B rotavirus)  
A;Reference number: A49810; MUID:94160571; PMID:8116245  
A;Accession: B49810  
A;Status: preliminary  
A;Molecule type: genomic RNA  
A;Residues: 1-320 <EID>  
A;Cross-references: GB:U01164; NID:G470378; PIDN:AAA60453.1; PID:G470379  
C;Superfamily: rotavirus B gene 7 protein

Query Match 61.9%; Score 39; DB 2; Length 320;  
Best Local Similarity 50.0%; Pred. No. 45;  
Matches 5; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GDVQDCIQM 10  
|:|:|:|:|:  
Db 157 GDKQCNCLRL 166

RESULT 8  
VGIVTH  
envelope glycoprotein - Thogoto virus (strain SiAr126)  
N;Alternate names: surface glycoprotein 75  
C;Species: Thogoto virus  
C;Date: 30-Sep-1992 #sequence\_revision 30-Sep-1992 #text\_change 16-Jul-1999  
C;Accession: A40821  
R;Morse, M.A.; Marriott, A.C.; Nuttall, P.A.  
Virology 186, 640-646, 1992  
A;Title: The glycoprotein of Thogoto virus (a tick-borne orthomyxo-like virus) is related  
A;Reference number: A40821; MUID:92124738; PMID:1733105  
A;Accession: A40821  
A;Molecule type: genomic RNA  
A;Residues: 1-512 <WOR>  
A;Cross-references: GB:M77280; NID:G335213; PIDN:AAA47918.1; PID:G335214  
C;Genetics:  
A;Map position: segment 4  
C;Superfamily: baculovirus major envelope glycoprotein  
C;Keywords: envelope protein; glycoprotein; transmembrane protein  
F;185,263,289,378,416/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 61.9%; Score 39; DB 1; Length 512;  
Best Local Similarity 71.4%; Pred. No. 67;  
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GDVQDC 7  
|:|:|:|:  
Db 109 GDICHDC 115

RESULT 9  
TI0567  
Probable serine/threonine-specific protein kinase (EC 2.7.1.1) - Arabidopsis th  
C;Species: Arabidopsis thaliana (mouse-ear cress)  
C;Date: 16-Jul-1999 #sequence\_revision 16-Jul-1999 #text\_change 21-Jan-2000  
C;Accession: TI0567  
R;Bevan, M.; Lennard, N.; Quail, M.; Harris, B.; Rajandream, M.A.; Barrell, B.G.; Bancroft  
submitted to the Protein Sequence Database, June 1999  
A;Reference number: Z16533  
A;Accession: TI0567  
A;Molecule type: DNA  
A;Residues: 1-666 <BEV>  
A;Cross-references: EMBL:AL050399; GSPDB:GN00062; ATSP:F25E4.90  
A;Experimental source: cultivar Columbia, BAC clone F25E4  
C;Genetics:  
A;Gene: ATSP:F25E4.90  
A;Map position: 4  
A;Introns: 315/1; 355/3; 426/1; 505/2; 556/3  
C;Superfamily: unassigned Ser/Thr or Tyr-specific protein kinases; protein kinase homolog  
C;Keywords: phosphotransferase; protein kinase  
F;337-618/Domain: protein kinase homology <KIN>

Query Match 61.9%; Score 39; DB 2; Length 666;

Best Local Similarity 75.0%; Pred. No. 83;  
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 3 VCQDCIQM 10  
Db 86 VCSDCIQL 93

RESULT 10  
T46918  
hypothetical protein DKFzp762L137.1 - human  
C:Species: Homo sapiens (man)  
C:Date: 17-Mar-2000 #sequence\_revision 17-Mar-2000 #text\_change 20-Apr-2000  
C:Accession: T46918  
R:Ottenwälder, B.; Obermaier, B.; Mewes, H.W.; Weil, B.; Wiemann, S.  
submitted to the Protein Sequence Database, February 2000  
A:Reference number: Z24136  
A:Accession: T46918  
A:Status: preliminary  
A:Molecule type: mRNA  
A:Residues: 1-221 <AAA>  
A:Cross-references: EMBL:AL157432  
A:Experimental source: adult melanoma (MeWo cell line); clone DKFzp762L137  
C:Genetics:  
A:Note: DKFzp762L137.1  
C:Superfamily: human hypothetical protein DKFzp762L137.1

Query Match 60.3%; Score 38; DB 2; Length 221;  
Best Local Similarity 45.5%; Pred. No. 48;  
Matches 5; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

Qy 1 GDVQCDCIQMV 11  
Db 50 GDICNACVLLV 60

RESULT 11  
T40043  
probable transport protein - fission yeast (Schizosaccharomyces pombe) (fragment)  
C:Species: Schizosaccharomyces pombe  
C:Date: 03-Dec-1999 #sequence\_revision 03-Dec-1999 #text\_change 03-Dec-1999  
C:Accession: T40043  
R:Wood, V.; Rajandream, M.A.; Barrell, B.G.; Skelton, J.; Churcher, C.M.  
submitted to the EMBL Data Library, September 1998  
A:Reference number: Z21901  
A:Accession: T40043  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 1-307 <WOO>  
A:Cross-references: EMBL:AL031517; PIDN:CAA20653.1; GSPDB:GN00067; SPDB:SPBC28E12.06c  
A:Experimental source: strain 972h-; cosmid c28E12  
C:Genetics:  
A:Gene: spbc3h7.16; SPDB:SPBC28E12.06c  
A:Map position: 2

Query Match 60.3%; Score 38; DB 2; Length 307;  
Best Local Similarity 50.0%; Pred. No. 63;  
Matches 5; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

Qy 2 DVCQDCIQMV 11  
Db 277 NVCSDCVSM 286

RESULT 12  
E70355  
conserved hypothetical protein aq\_622 - Aquifex aeolicus  
C:Species: Aquifex aeolicus  
C:Date: 08-May-1998 #sequence\_revision 08-May-1998 #text\_change 05-Nov-1999  
C:Accession: E70355  
R:Deckert, G.; Warren, P.V.; Gaasterland, T.; Young, W.G.; Lenox, A.L.; Graham, D.E.; Ov  
V.  
Nature 392, 353-358, 1998

A:Title: The complete genome of the hyperthermophilic bacterium Aquifex aeolicus.  
A:Reference number: A70300; MUID:98196666; PMID:9537320  
A:Accession: E70355  
A:Status: preliminary; nucleic acid sequence not shown; translation not shown  
A:Molecule type: DNA  
A:Residues: 1-320 <AQF>  
A:Cross-references: GB:AE000698; NID:g2983224; PIDN:AAC06834.1; PID:g2983237; GB:AE00065  
A:Experimental source: strain VF5  
C:Genetics:  
A:Gene: aq\_622

Query Match 60.3%; Score 38; DB 2; Length 320;  
Best Local Similarity 60.0%; Pred. No. 65;  
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Qy 2 DVCQDCIQMV 11  
Db 179 EVCLDCIDML 188

RESULT 13  
A57234  
lin-44 protein precursor - Caenorhabditis elegans  
C:Species: Caenorhabditis elegans  
C:Date: 23-Feb-1996 #sequence\_revision 23-Feb-1996 #text\_change 20-Sep-1999  
C:Accession: A57234  
R:Herman, M.A.; Vassilieva, L.L.; Horvitz, H.R.; Shaw, J.E.; Herman, R.K.  
Cell 83, 101-110, 1995  
A:Title: The Caenorhabditis elegans gene lin-44, which controls the polarity of certain  
submitted to the EMBL Data Library, September 1995  
A:Reference number: A57234; MUID:96006529; PMID:7553861  
A:Accession: A57234  
A:Status: preliminary  
A:Molecule type: mRNA  
A:Residues: 1-348 <HER>  
A:Cross-references: GB:U22179; NID:g758776; PID:g758777  
C:Superfamily: int-1 transforming protein

Query Match 60.3%; Score 38; DB 2; Length 348;  
Best Local Similarity 85.7%; Pred. No. 70;  
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 3 VCQDCIQ 9  
Db 325 VCKDCIQ 331

RESULT 14  
S41686  
geranylgeranyltransferase type I (EC 2.5.1.-) beta chain - fission yeast (Schizosaccharo  
C:Species: Schizosaccharomyces pombe  
C:Date: 31-Mar-1992 #sequence\_revision 14-Sep-1994 #text\_change 21-Jan-2000  
C:Accession: S41686  
R:Diaz, M.; Sanchez, Y.; Bennett, T.; Sun, C.R.; Godoy, C.; Tamanoi, F.; Duran, A.; Pere  
EMBO J. 12, 5245-5254, 1993  
A:Title: The Schizosaccharomyces pombe cwg2(+) gene codes for the beta subunit of a gera  
A:Reference number: S41686; MUID:94085400; PMID:8262067  
A:Accession: S41686  
A:Molecule type: DNA  
A:Residues: 1-355 <DIA>  
A:Cross-references: EMBL:Z12155; NID:g4946; PIDN:CAA78143.1; PID:g396477  
C:Genetics:  
A:Gene: cwg2  
C:Superfamily: cell division control protein CDC43  
C:Keywords: glycoprotein; transferase; transmembrane protein  
F:31-47/Domain: transmembrane #status predicted <TM1>  
F:95-111/Domain: transmembrane #status predicted <TM2>  
F:113,215,219,222/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 60.3%; Score 38; DB 2; Length 355;  
Best Local Similarity 62.5%; Pred. No. 71;  
Matches 5; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

Qy 2 DVCQDCIQ 9

```

Db          343 NICKDCIQ 350
:|:|:|
RESULT 15
T24442
hypothetical protein T04B2.5 - Caenorhabditis elegans
C;Species: Caenorhabditis elegans
C;Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 21-Apr-2003
C;Accession: T24442; T26042
R;Coles, L.
submitted to the EMBL Data Library, December 1995
A;Reference number: Z19890
A;Accession: T24442
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-456 <WIL>
A;Cross-references: EMBL:Z68299; PIDN:CAA92615.1; GSPDB:GN00022; CESP:T04B2.5
A;Experimental source: clone T04B2
R;Coles, L.
submitted to the EMBL Data Library, December 1995
A;Reference number: Z20141
A;Accession: T26042
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-456 <W12>
A;Cross-references: EMBL:Z68301; PIDN:CAA92628.1; GSPDB:GN00022; CESP:T04B2.5
A;Experimental source: clone W01B6
C;Genetics:
A;Gene: CESP:T04B2.5
A;Map position: 4
A;Introns: 48/3; 104/3; 139/1; 424/3
C;Superfamily: patatin

Query Match      60.3%; Score 38; DB 2; Length 456;
Best Local Similarity 50.0%; Pred. No. 88;
Matches 5; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY      1 GDVCCDCIQ 10
Db      97 GSVCECAEM 106
| | | | |
| | | | |

Search completed: May 5, 2004, 13:34:35
Job time : 0.438746 secs

```

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 5, 2004, 13:14:53 ; Search time 1.16999 Seconds  
(without alignments)  
2606.077 Million cell updates/sec

Title: US-09-743-684A-19  
Perfect score: 63  
Sequence: 1 GDVQCDCIQMV 11

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1138120 seqs, 277189581 residues

Total number of hits satisfying chosen parameters: 1138120

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA:\*

- 1: /cgn2\_6/ptodata/2/pubpaa/US07\_PUBCOMB.pep.\*
- 2: /cgn2\_6/ptodata/2/pubpaa/PCT\_NEW\_PUB.pep.\*
- 3: /cgn2\_6/ptodata/2/pubpaa/US06\_NEW\_PUB.pep.\*
- 4: /cgn2\_6/ptodata/2/pubpaa/US06\_PUBCOMB.pep.\*
- 5: /cgn2\_6/ptodata/2/pubpaa/US07\_NEW\_PUB.pep.\*
- 6: /cgn2\_6/ptodata/2/pubpaa/PCTUS\_PUBCOMB.pep.\*
- 7: /cgn2\_6/ptodata/2/pubpaa/US08\_NEW\_PUB.pep.\*
- 8: /cgn2\_6/ptodata/2/pubpaa/US08\_PUBCOMB.pep.\*
- 9: /cgn2\_6/ptodata/2/pubpaa/US09A\_PUBCOMB.pep.\*
- 10: /cgn2\_6/ptodata/2/pubpaa/US09B\_PUBCOMB.pep.\*
- 11: /cgn2\_6/ptodata/2/pubpaa/US09C\_PUBCOMB.pep.\*
- 12: /cgn2\_6/ptodata/2/pubpaa/US09\_NEW\_PUB.pep.\*
- 13: /cgn2\_6/ptodata/2/pubpaa/US10A\_PUBCOMB.pep.\*
- 14: /cgn2\_6/ptodata/2/pubpaa/US10B\_PUBCOMB.pep.\*
- 15: /cgn2\_6/ptodata/2/pubpaa/US10C\_PUBCOMB.pep.\*
- 16: /cgn2\_6/ptodata/2/pubpaa/US10\_NEW\_PUB.pep.\*
- 17: /cgn2\_6/ptodata/2/pubpaa/US60\_NEW\_PUB.pep.\*
- 18: /cgn2\_6/ptodata/2/pubpaa/US60\_PUBCOMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	63	100.0	209	14	US-10-043-487-340
2	63	100.0	479	10	US-09-978-418-40
3	63	100.0	523	9	US-09-767-007A-2
4	63	100.0	524	9	US-09-870-759-60
5	63	100.0	524	10	US-09-751-708A-60
6	63	100.0	524	12	US-10-267-502-386
7	63	100.0	527	9	US-09-870-759-61
8	63	100.0	527	10	US-09-751-708A-61
9	63	100.0	527	14	US-10-060-036-73
10	57	90.5	156	12	US-09-925-298-644
11	57	90.5	156	14	US-10-102-806-644
12	47	74.6	554	14	US-10-205-194-176
13	42	66.7	521	15	US-10-276-162-1
14	42	66.7	531	11	US-09-833-245-903
15	41	65.1	60	14	US-10-083-357-929

16	40	63.5	22	14	US-10-300-694A-73	Sequence 73, Appl
17	40	63.5	532	14	US-10-128-714-3442	Sequence 3442, Ap
18	40	63.5	532	14	US-10-128-714-8442	Sequence 8442, Ap
19	39	61.9	48	12	US-10-424-599-148377	Sequence 148377,
20	39	61.9	171	12	US-10-424-599-269800	Sequence 269800,
21	39	61.9	362	12	US-10-332-426-8	Sequence 8, Appli
22	39	61.9	653	14	US-10-081-872-194	Sequence 194, App
23	39	61.9	653	15	US-10-385-303-194	Sequence 194, App
24	38	60.3	166	12	US-10-425-114-68084	Sequence 68084, A
25	38	60.3	221	12	US-10-170-385-441	Sequence 441, App
26	38	60.3	251	14	US-10-106-698-4737	Sequence 4737, Ap
27	38	60.3	252	9	US-09-925-301-993	Sequence 993, App
28	38	60.3	469	12	US-10-282-122A-50094	Sequence 50094, A
29	37	58.7	110	9	US-09-764-864-848	Sequence 848, App
30	37	58.7	129	12	US-10-424-599-249703	Sequence 249703,
31	37	58.7	166	12	US-10-424-599-277844	Sequence 277844,
32	37	58.7	178	9	US-09-764-864-1303	Sequence 1303, Ap
33	37	58.7	319	9	US-09-764-868-659	Sequence 659, App
34	37	58.7	333	12	US-10-243-552-875	Sequence 875, App
35	37	58.7	365	13	US-10-029-180-62	Sequence 62, Appl
36	37	58.7	429	15	US-10-369-493-22639	Sequence 22639, A
37	37	58.7	465	14	US-10-103-313-357	Sequence 357, App
38	37	58.7	530	12	US-10-282-122A-74259	Sequence 74259, A
39	37	58.7	595	15	US-10-369-493-11077	Sequence 11077, A
40	37	58.7	748	15	US-10-369-493-18341	Sequence 18341, A
41	37	58.7	760	12	US-10-282-122A-74644	Sequence 74644, A
42	37	58.7	766	15	US-10-094-749-3040	Sequence 3040, Ap
43	37	58.7	766	15	US-10-231-913-99	Sequence 99, Appl
44	37	58.7	766	15	US-10-231-913-101	Sequence 101, App
45	37	58.7	767	15	US-10-262-445-36	Sequence 36, Appl

ALIGNMENTS

RESULT 1

US-10-043-487-340  
; Sequence 340, Application US/10043487  
; Publication No. US20030055220A1  
; GENERAL INFORMATION:  
; APPLICANT: HYBRIGENICS  
; APPLICANT: Pierre, LEGRAIN  
; TITLE OF INVENTION: Protein-protein interactions between Shigella Flexneri polypeptides  
; TITLE OF INVENTION: mammalian polypeptides  
; FILE REFERENCE: B4778A  
; CURRENT APPLICATION NUMBER: US/10/043,487  
; CURRENT FILING DATE: 2002-04-30  
; PRIOR FILING DATE: 2001-01-12  
; PRIOR FILING DATE: 2001-01-12  
; NUMBER OF SEQ ID NOS: 561  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 340  
; LENGTH: 209  
; TYPE: PRT  
; ORGANISM: Shigella Flexneri  
US-10-043-487-340

Query Match 100.0%; Score 63; DB 14; Length 209;  
Best Local Similarity 100.0%; Pred. No. 0.015;  
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qv 1 GDVQCDCIQMV 11

Db 45 GDVQCDCIQMV 55

RESULT 2

US-09-978-418-40  
; Sequence 40, Application US/09978418  
; Publication No. US20030118997A1  
; GENERAL INFORMATION:  
; APPLICANT: Bejanin, Stephan  
; APPLICANT: Tanaka, Hiroaki

```
; TITLE OF INVENTION: HUMAN CDNAS AND PROTEINS AND USES THEREOF
; FILE REFERENCE: 142 US5 REG
; CURRENT APPLICATION NUMBER: US/09/978,418
; CURRENT FILING DATE: 2002-10-15
; PRIOR APPLICATION NUMBER: 60/311,305
; PRIOR FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: 60/314,734
; PRIOR FILING DATE: 2001-08-24
; PRIOR APPLICATION NUMBER: 60/318,204
; PRIOR FILING DATE: 2001-09-07
; PRIOR APPLICATION NUMBER: 60/326,470
; PRIOR FILING DATE: 2001-10-01
; NUMBER OF SEQ ID NOS: 52
; SOFTWARE: JPatent
; SEQ ID NO 40
; LENGTH: 479
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-978-418-40

Query Match      100.0%; Score 63; DB 10; Length 479;
Best Local Similarity 100.0%; Pred. No. 0.034;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 GDVQCDCIQMV 11
Db      195 GDVQCDCIQMV 205

RESULT 3
US-09-767-007A-2
; Sequence 2, Application US/09767007A
; Patent No. US2002007725A1
; GENERAL INFORMATION:
; APPLICANT: John S. O'Brien
; APPLICANT: Yasuo Kishimoto
; TITLE OF INVENTION: PHARMACEUTICAL COMPOSITIONS COMPRISING
; FILE REFERENCE: MYELOS.2DC1C1
; CURRENT APPLICATION NUMBER: US/09/767,007A
; CURRENT FILING DATE: 2001-01-22
; PRIOR APPLICATION NUMBER: 08/958,970
; PRIOR FILING DATE: 1997-10-28
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 523
; TYPE: PRT
; ORGANISM: homo sapiens
US-09-767-007A-2

Query Match      100.0%; Score 63; DB 9; Length 523;
Best Local Similarity 100.0%; Pred. No. 0.037;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 GDVQCDCIQMV 11
Db      194 GDVQCDCIQMV 204

RESULT 4
US-09-870-759-60
; Sequence 60, Application US/09870759
; Patent No. US20020177551A1
; GENERAL INFORMATION:
; APPLICANT: TERMAN, David S
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATMENT OF NEOPLASTIC DISEASE
; FILE REFERENCE: 870759
; CURRENT APPLICATION NUMBER: US/09/870,759
; CURRENT FILING DATE: 2002-01-14
; PRIOR APPLICATION NUMBER: US 60/208,128
; PRIOR FILING DATE: 2000-05-30
; NUMBER OF SEQ ID NOS: 166

; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 60
; LENGTH: 524
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-870-759-60

Query Match      100.0%; Score 63; DB 9; Length 524;
Best Local Similarity 100.0%; Pred. No. 0.037;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 GDVQCDCIQMV 11
Db      195 GDVQCDCIQMV 205

RESULT 5
US-09-751-708A-60
; Sequence 60, Application US/09751708A
; Publication No. US20030157113A1
; GENERAL INFORMATION:
; APPLICANT: TERMAN, David S
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATMENT OF NEOPLASTIC DISEASE
; FILE REFERENCE: 751708
; CURRENT APPLICATION NUMBER: US/09/751,708A
; CURRENT FILING DATE: 2002-10-15
; PRIOR APPLICATION NUMBER: US 60/173,371
; PRIOR FILING DATE: 1999-12-28
; NUMBER OF SEQ ID NOS: 166
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 60
; LENGTH: 524
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-751-708A-60

Query Match      100.0%; Score 63; DB 10; Length 524;
Best Local Similarity 100.0%; Pred. No. 0.037;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 GDVQCDCIQMV 11
Db      195 GDVQCDCIQMV 205

RESULT 6
US-10-267-502-386
; Sequence 386, Application US/10267502
; Publication No. US20040071700A1
; GENERAL INFORMATION:
; APPLICANT: Kim, Jaeseob
; APPLICANT: Galant, Ron
; TITLE OF INVENTION: Obesity Linked Genes
; FILE REFERENCE: LSD-07416
; CURRENT APPLICATION NUMBER: US/10/267,502
; CURRENT FILING DATE: 2003-01-27
; NUMBER OF SEQ ID NOS: 439
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 386
; LENGTH: 524
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-267-502-386

Query Match      100.0%; Score 63; DB 12; Length 524;
Best Local Similarity 100.0%; Pred. No. 0.037;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 GDVQCDCIQMV 11
Db      195 GDVQCDCIQMV 205
```

```
RESULT 7
US-09-870-759-61
; Sequence 61, Application US/09870759
; Patent No. US20020177551A1
; GENERAL INFORMATION:
; APPLICANT: TERMAN, David S
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATMENT OF NEOPLASTIC DISEASE
; FILE REFERENCE: 870759
; CURRENT APPLICATION NUMBER: US/09/870,759
; CURRENT FILING DATE: 2002-01-14
; PRIOR APPLICATION NUMBER: US 60/208,128
; PRIOR FILING DATE: 2000-05-30
; NUMBER OF SEQ ID NOS: 166
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 61
; LENGTH: 527
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-870-759-61

Query Match      100.0%; Score 63; DB 9; Length 527;
Best Local Similarity 100.0%; Pred. No. 0.037;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GDVCQDCIQMV 11
Db 195 GDVCQDCIQMV 205

RESULT 8
US-09-751-708A-61
; Sequence 61, Application US/09751708A
; Publication No. US20030157113A1
; GENERAL INFORMATION:
; APPLICANT: TERMAN, David S
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATMENT OF NEOPLASTIC DISEASE
; FILE REFERENCE: 751708
; CURRENT APPLICATION NUMBER: US/09/751,708A
; CURRENT FILING DATE: 2002-10-15
; PRIOR APPLICATION NUMBER: US 60/173,371
; PRIOR FILING DATE: 1999-12-28
; NUMBER OF SEQ ID NOS: 165
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 61
; LENGTH: 527
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-751-708A-61

Query Match      100.0%; Score 63; DB 10; Length 527;
Best Local Similarity 100.0%; Pred. No. 0.037;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GDVCQDCIQMV 11
Db 195 GDVCQDCIQMV 205

RESULT 9
US-10-060-036-73
; Sequence 73, Application US/10060036
; Publication No. US20030073144A1
; GENERAL INFORMATION:
; APPLICANT: Benson, Darin R.
; APPLICANT: Kalos, Michael D.
; APPLICANT: Lodes, Michael J.
; APPLICANT: Persing, David H.
; APPLICANT: Hepler, William T.
; APPLICANT: Jiang, Yugu
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; TITLE OF INVENTION: AND DIAGNOSIS OF PANCREATIC CANCER
; FILE REFERENCE: 210121.566
; CURRENT APPLICATION NUMBER: US/10/060,036
```

```
; CURRENT FILING DATE: 2002-01-30
; NUMBER OF SEQ ID NOS: 4560
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 73
; LENGTH: 527
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-060-036-73

Query Match      100.0%; Score 63; DB 14; Length 527;
Best Local Similarity 100.0%; Pred. No. 0.037;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GDVCQDCIQMV 11
Db 195 GDVCQDCIQMV 205

RESULT 10
US-09-925-298-644
; Sequence 644, Application US/09925298
; Publication No. US20020039764A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
; FILE REFERENCE: PA103
; CURRENT APPLICATION NUMBER: US/09/925,298
; CURRENT FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: PCT/US00/05881
; PRIOR FILING DATE: 2000-03-08
; PRIOR APPLICATION NUMBER: 60/124,270
; PRIOR FILING DATE: 1999-03-12
; NUMBER OF SEQ ID NOS: 846
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 644
; LENGTH: 156
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (12)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-09-925-298-644

Query Match      90.5%; Score 57; DB 12; Length 156;
Best Local Similarity 100.0%; Pred. No. 0.11;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 DVCQDCIQMV 11
Db 13 DVCQDCIQMV 22

RESULT 11
US-10-102-806-644
; Sequence 644, Application US/10102806
; Publication No. US20030054421A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
; FILE REFERENCE: PA103P1C1
; CURRENT APPLICATION NUMBER: US/10/102,806
; CURRENT FILING DATE: 2002-03-22
; PRIOR APPLICATION NUMBER: 09/925,298
; PRIOR FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: PCT/US00/05881
; PRIOR FILING DATE: 2000-03-08
; PRIOR APPLICATION NUMBER: 60/124,270
; PRIOR FILING DATE: 1999-03-12
; NUMBER OF SEQ ID NOS: 846
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 644
; LENGTH: 156
```

```

; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (12)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-10-102-806-644

Query Match          90.5%; Score 57; DB 14; Length 156;
Best Local Similarity 100.0%; Pred. No. 0.11;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy  2 DVCQDCIQMV 11
    |||||
Db  13 DVCQDCIQMV 22

RESULT 12
US-10-205-194-176
; Sequence 176, Application US/10205194
; Publication No. US20030134301A1
; GENERAL INFORMATION:
; APPLICANT: Warner-Lambert Company
; APPLICANT: Lee, Kevin
; APPLICANT: Dixon, Alistair
; APPLICANT: Brooksbank, Robert
; APPLICANT: Pincock, Robert
; TITLE OF INVENTION: Identification and Use of Molecules Implicated in Pain
; FILE REFERENCE: WL-A-018201
; CURRENT APPLICATION NUMBER: US/10/205,194
; PRIOR FILING DATE: 2000-07-24
; PRIOR APPLICATION NUMBER: GB 0118354.0
; PRIOR FILING DATE: 2001-07-27
; NUMBER OF SEQ ID NOS: 177
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 176
; LENGTH: 554
; TYPE: PRT
; ORGANISM: Rattus sp.
; FEATURE:
; OTHER INFORMATION: Prosaposin
US-10-205-194-176

Query Match          74.6%; Score 47; DB 14; Length 554;
Best Local Similarity 70.0%; Pred. No. 15;
Matches 7; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

Qy  2 DVCQDCIQMV 11
    |||||
Db  195 DVCQDCMKLV 204

RESULT 13
US-10-276-162-1
; Sequence 1, Application US/10276162
; Publication No. US20030215822A1
; GENERAL INFORMATION:
; APPLICANT: GRIFFIN, Jennifer A.
; APPLICANT: YAO, Monique G.
; APPLICANT: BRUNS, Christopher M.
; APPLICANT: YUE, Henry
; APPLICANT: DELEGEANE, Angelo M.
; APPLICANT: HAPALIA, April
; APPLICANT: PATTERSON, Chandra
; APPLICANT: POLICKY, Jennifer L.
; APPLICANT: TRIBOULEY, Catherine M.
; APPLICANT: BAUGHN, Mariah R.
; APPLICANT: NGUYEN, Daniel B.
; APPLICANT: LAL, Preeti
; APPLICANT: TANG, Y. Tom
; APPLICANT: JACKSON, Jennifer L.
; APPLICANT: LU, Dyung Aina M.
; APPLICANT: BATRA, Sajeev

```

```

; APPLICANT: AU-YOUNG, Janice
; APPLICANT: REDDY, Roopa
; APPLICANT: AZIMZAI, Yalda
; TITLE OF INVENTION: SECRETED PROTEINS
; FILE REFERENCE: PI-0071 USN
; CURRENT APPLICATION NUMBER: US/10/276,162
; CURRENT FILING DATE: 2002-10-15
; PRIOR APPLICATION NUMBER: US01/11861
; PRIOR FILING DATE: 2001-04-11
; PRIOR APPLICATION NUMBER: 60/197,854
; PRIOR FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/202,373
; PRIOR FILING DATE: 2000-05-04
; PRIOR APPLICATION NUMBER: 60/205,899
; PRIOR FILING DATE: 2000-05-18
; PRIOR APPLICATION NUMBER: 60/210,155
; PRIOR FILING DATE: 2000-06-01
; PRIOR APPLICATION NUMBER: 60/209,401
; PRIOR FILING DATE: 2000-06-01
; NUMBER OF SEQ ID NOS: 28
; SOFTWARE: PERL Program
; SEQ ID NO 1
; LENGTH: 521
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: Incyte ID No. US20030215822A1 7473577CDI
US-10-276-162-1

```

```

Query Match          66.7%; Score 42; DB 15; Length 521;
Best Local Similarity 54.5%; Pred. No. 91;
Matches 6; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

Qy  1 GDVCQDCIQMV 11
    |::|::|
Db  181 GALTQDCVQV 191

```

```

RESULT 14
US-09-833-245-903
; Sequence 903, Application US/09833245
; Publication No. US20040010134A1
; GENERAL INFORMATION:
; APPLICANT: Human Genome Sciences, Inc.
; TITLE OF INVENTION: Albumin Fusion Proteins
; FILE REFERENCE: PF546PCT
; CURRENT APPLICATION NUMBER: US/09/833,245
; CURRENT FILING DATE: 2001-04-12
; PRIOR APPLICATION NUMBER: 60/229, 358
; PRIOR FILING DATE: 2000-04-12
; PRIOR APPLICATION NUMBER: 60/256, 931
; PRIOR FILING DATE: 2000-12-21
; PRIOR APPLICATION NUMBER: 60/199, 384
; PRIOR FILING DATE: 2000-04-25
; NUMBER OF SEQ ID NOS: 2267
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 903
; LENGTH: 531
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-833-245-903

```

```

Query Match          66.7%; Score 42; DB 11; Length 531;
Best Local Similarity 54.5%; Pred. No. 92;
Matches 6; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

```

```

Qy  1 GDVCQDCIQMV 11
    |::|::|
Db  181 GALTQDCVQV 191

```

RESULT 15



```
US-10-083-357-929
; Sequence 929, Application US/10083357
; Publication No. US20030054370A1
; GENERAL INFORMATION:
; APPLICANT: Qiandong Zeng et al.
; TITLE OF INVENTION: Systemic Discovery of New Genes
; FILE REFERENCE: 032796-090
; CURRENT APPLICATION NUMBER: US/10/083,357
; CURRENT FILING DATE: 2002-02-27
; NUMBER OF SEQ ID NOS: 1346
; SEQ ID NO 929
; LENGTH: 60
; TYPE: PRT
; ORGANISM: Saccharomyces cerevisiae
US-10-083-357-929

Query Match      65.1%; Score 41; DB 14; Length 60;
Best Local Similarity 66.7%; Pred. No. 16;
Matches 6; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy      1 GDVCQDCIQ 9
        |||||
Db      28 GDACSDCFQ 36

Search completed: May 5, 2004, 13:28:24
Job time : 1.16999 secs
```

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 5, 2004, 13:10:37 ; Search time 0.480532 Seconds  
(without alignments)  
1181.786 Million cell updates/sec

Title: US-09-743-684A-19  
Perfect score: 63  
Sequence: 1 GDVQCDCIQMV 11

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 389414 seqs, 51625971 residues

Total number of hits satisfying chosen parameters: 389414

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Issued Patents AA.\*  
1: /cgn2\_6/ptodata/2/iaa/5A COMB.pep.\*  
2: /cgn2\_6/ptodata/2/iaa/5B COMB.pep.\*  
3: /cgn2\_6/ptodata/2/iaa/6A COMB.pep.\*  
4: /cgn2\_6/ptodata/2/iaa/6B COMB.pep.\*  
5: /cgn2\_6/ptodata/2/iaa/PCTUS COMB.pep.\*  
6: /cgn2\_6/ptodata/2/iaa/backfiles.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	63	100.0	11	4	US-09-352-548-19
2	63	100.0	80	2	US-08-584-671-15
3	63	100.0	80	3	US-09-027-376-15
4	63	100.0	80	3	US-09-094-192-15
5	63	100.0	81	4	US-09-352-548-2
6	63	100.0	523	1	US-08-100-247-2
7	63	100.0	523	1	US-08-483-146A-2
8	63	100.0	523	1	US-08-232-513A-3
9	63	100.0	523	1	US-08-484-594A-2
10	63	100.0	523	4	US-09-076-258A-2
11	63	100.0	523	4	US-08-756-031-2
12	63	100.0	524	4	US-09-352-548-1
13	57	90.5	10	4	US-09-352-548-21
14	57	90.5	16	4	US-09-352-548-48
15	55	87.3	13	5	US-09-352-548-13
16	53	84.1	9	4	US-09-352-548-24
17	53	84.1	11	4	US-09-352-548-32
18	53	84.1	11	4	US-09-352-548-33
19	51	81.0	9	4	US-09-352-548-29
20	48	76.2	8	4	US-09-352-548-25
21	47	74.6	8	4	US-09-352-548-30
22	47	74.6	19	2	US-08-584-671-5
23	47	74.6	19	3	US-09-027-376-5
24	47	74.6	19	3	US-09-094-192-5
25	47	74.6	79	2	US-08-584-671-14
26	47	74.6	79	2	US-08-584-671-16
27	47	74.6	79	3	US-09-027-376-14

28	47	74.6	79	3	US-09-027-376-16	Sequence 16, Appl
29	47	74.6	79	3	US-09-094-192-14	Sequence 14, Appl
30	47	74.6	79	3	US-09-094-192-16	Sequence 16, Appl
31	43	68.3	7	4	US-09-352-548-26	Sequence 26, Appl
32	43	68.3	11	4	US-09-352-548-20	Sequence 20, Appl
33	40	63.5	15	2	US-08-584-671-4	Sequence 4, Appl
34	40	63.5	15	3	US-09-027-376-4	Sequence 4, Appl
35	40	63.5	15	3	US-09-094-192-4	Sequence 4, Appl
36	39	61.9	6	4	US-09-352-548-27	Sequence 27, Appl
37	39	61.9	16	4	US-09-352-548-55	Sequence 55, Appl
38	38	60.3	7	4	US-09-352-548-31	Sequence 31, Appl
39	36	57.1	103	2	US-09-109-266-4	Sequence 4, Appl
40	36	57.1	252	1	US-08-411-777-8	Sequence 8, Appl
41	36	57.1	252	3	US-09-057-088-8	Sequence 8, Appl
42	36	57.1	341	3	US-09-008-465-1	Sequence 1, Appl
43	36	57.1	341	4	US-09-528-959-1	Sequence 1, Appl
44	35	55.6	52	4	US-09-621-976-6386	Sequence 6386, Ap
45	35	55.6	72	3	US-08-338-579A-101	Sequence 101, App

## ALIGNMENTS

## RESULT 1

US-09-352-548-19  
; Sequence 19, Application US/09352548  
; Patent No. 6500431  
; GENERAL INFORMATION:  
; APPLICANT: Gill, Parkash S.  
; APPLICANT: Parkash S. Gill, M.D., Inc.  
; TITLE OF INVENTION: No. 6500431el Inhibitors of Angiogenesis and Tumor Growth  
; FILE REFERENCE: 017986-000410US  
; CURRENT APPLICATION NUMBER: US/09/352,548  
; CURRENT FILING DATE: 1999-07-12  
; EARLIER APPLICATION NUMBER: US 60/092,647  
; EARLIER FILING DATE: 1998-07-13  
; NUMBER OF SEQ ID NOS: 59  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 19  
; LENGTH: 11  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence:G1-V11  
; OTHER INFORMATION: anti-angiogenic polypeptide  
US-09-352-548-19

Query Match 100.0%; Score 63; DB 4; Length 11;  
Best Local Similarity 100.0%; Pred. No. 0.00033;  
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GDVQCDCIQMV 11  
| | | | | | | | | |  
Db 1 GDVQCDCIQMV 11

## RESULT 2

US-08-584-671-15  
; Sequence 15, Application US/08584671  
; Patent No. 5910568  
; GENERAL INFORMATION:  
; APPLICANT: HAMMERSTEDT, ROY H, BARBATO, GUY F,  
; APPLICANT: CRAMER, PALMER  
; TITLE OF INVENTION: MOLECULE INVOLVED IN BINDING OF SPERM  
; TITLE OF INVENTION: TO EGG SURFACES AND PROCEDURES FOR USE OF THIS MOLECULE  
; NUMBER OF SEQUENCES: 16  
; CORRESPONDENCE ADDRESSES:  
; ADDRESSEE: INTELLECTUAL PROPERTY OFFICE, THE PENNSYLVANIA  
; ADDRESSEE: STATE UNIVERSITY  
; STREET: 113 TECHNOLOGY CENTER  
; CITY: UNIVERSITY PARK  
; STATE: PENNSYLVANIA

COUNTRY: UNITED STATES OF AMERICA  
ZIP: 16802-7000  
COMPUTER READABLE FORM:  
MEDIUM TYPE: FLOPPY DISK  
COMPUTER: NEC 286  
OPERATING SYSTEM: DOS  
SOFTWARE: WORDPERFECT 5.1  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/584,671  
FILING DATE:  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: MONAHAN, THOMAS J  
REGISTRATION NUMBER: 29835  
REFERENCE/DOCKET NUMBER:  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 814-865-6277  
TELEFAX: 814-865-3591  
INFORMATION FOR SEQ ID NO: 15:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 80  
TYPE: AMINO ACID  
STRANDEDNESS: SINGLE  
TOPOLOGY: UNKNOWN  
US-08-584-671-15

Query Match 100.0%; Score 63; DB 2; Length 80;  
Best Local Similarity 100.0%; Pred. No. 0.0021;  
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GDVQCDCIQMV 11  
Db 67 GDVQCDCIQMV 77

## RESULT 3

US-09-027-376-15  
Sequence 15, Application US/09027376  
Patent No. 6004586  
GENERAL INFORMATION:  
APPLICANT: HAMMERSTEDT, ROY H, BARBATO, GUY F,  
APPLICANT: CRAWER, PALMER  
TITLE OF INVENTION: MOLECULE INVOLVED IN BINDING OF SPERM  
TITLE OF INVENTION: TO EGG SURFACES AND PROCEDURES FOR USE OF THIS MOLECULE  
TITLE OF INVENTION: TO ENHANCE OR DECREASE POTENTIAL FERTILITY  
NUMBER OF SEQUENCES: 16  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: INTELLECTUAL PROPERTY OFFICE, THE PENNSYLVANIA  
ADDRESSEE: STATE UNIVERSITY  
STREET: 113 TECHNOLOGY CENTER  
CITY: UNIVERSITY PARK  
STATE: PENNSYLVANIA  
COUNTRY: UNITED STATES OF AMERICA  
ZIP: 16802-7000  
COMPUTER READABLE FORM:  
MEDIUM TYPE: FLOPPY DISK  
COMPUTER: NEC 286  
OPERATING SYSTEM: DOS  
SOFTWARE: WORDPERFECT 5.1  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/027,376  
FILING DATE:  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/584,671  
FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: MONAHAN, THOMAS J  
REGISTRATION NUMBER: 29835  
REFERENCE/DOCKET NUMBER:  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 814-865-6277  
TELEFAX: 814-865-3591

INFORMATION FOR SEQ ID NO: 15:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 80  
TYPE: AMINO ACID  
STRANDEDNESS: SINGLE  
TOPOLOGY: UNKNOWN  
US-09-027-376-15

Query Match 100.0%; Score 63; DB 3; Length 80;  
Best Local Similarity 100.0%; Pred. No. 0.0021;  
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GDVQCDCIQMV 11  
Db 67 GDVQCDCIQMV 77

## RESULT 4

US-09-094-192-15  
Sequence 15, Application US/09094192  
Patent No. 6103483  
GENERAL INFORMATION:  
APPLICANT: HAMMERSTEDT, ROY H., BARBATO, GUY F.  
APPLICANT: CRAWER, PALMER  
TITLE OF INVENTION: MOLECULE INVOLVED IN BINDING OF SPERM TO EGG SURFACES AND PROC  
NUMBER OF SEQUENCES: 16  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: INTELLECTUAL PROPERTY OFFICE, THE PENNSYLVANIA STATE UNIVERSITY  
STREET: 113 TECHNOLOGY CENTER  
CITY: UNIVERSITY PARK  
STATE: PENNSYLVANIA  
COUNTRY: UNITED STATES OF AMERICA  
ZIP: 16802-7000  
COMPUTER READABLE FORM:  
MEDIUM TYPE: FLOPPY DISK  
COMPUTER: NEC 286  
OPERATING SYSTEM: DOS  
SOFTWARE: WORDPERFECT 5.1  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/094,192  
FILING DATE:  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: MONAHAN, THOMAS J  
REGISTRATION NUMBER: 29835  
REFERENCE/DOCKET NUMBER:  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 814-865-6277  
TELEFAX: 814-865-3591  
INFORMATION FOR SEQ ID NO: 15:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 80  
TYPE: AMINO ACID  
STRANDEDNESS: SINGLE  
TOPOLOGY: UNKNOWN  
US-09-094-192-15

Query Match 100.0%; Score 63; DB 3; Length 80;  
Best Local Similarity 100.0%; Pred. No. 0.0021;  
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GDVQCDCIQMV 11  
Db 67 GDVQCDCIQMV 77

## RESULT 5

US-09-352-548-2  
Sequence 2, Application US/09352548  
Patent No. 6500431  
GENERAL INFORMATION:  
APPLICANT: Gill, Parkash S.  
APPLICANT: Parkash S. Gill, M.D., Inc.

;; TITLE OF INVENTION: No. 6500431a1 Inhibitors of Angiogenesis and Tumor Growth

;; FILE REFERENCE: 017986-000410US

;; CURRENT APPLICATION NUMBER: US/09/352,548

;; CURRENT FILING DATE: 1999-07-12

;; EARLIER APPLICATION NUMBER: US 60/092,647

;; EARLIER FILING DATE: 1998-07-13

;; NUMBER OF SEQ ID NOS: 59

;; SOFTWARE: Patent in Ver. 2.1

;; SEQ ID NO 2

;; LENGTH: 81

;; TYPE: PRT

;; ORGANISM: Homo sapiens

;; FEATURE:

;; OTHER INFORMATION: Saposin B

;; US-09-352-548-2

Query Match 100.0%; Score 63; DB 4; Length 81;

Best Local Similarity 100.0%; Pred. No. 0.0021;

Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GDVCQDCIQMV 11

Db 1 GDVCQDCIQMV 11

#### RESULT 6

US-08-100-247-2

;; Sequence 2, Application US/08100247

;; Patent No. 5571787

;; GENERAL INFORMATION:

;; APPLICANT: O'BRIEN, JOHN S.

;; APPLICANT: KISHIMOTO, YASUO

;; TITLE OF INVENTION: PROSAPOSIN AS A NEUROTROPHIC FACTOR

;; NUMBER OF SEQUENCES: 5

;; CORRESPONDENCE ADDRESS:

;; ADDRESSEE: KNOBBE, MARTENS, OLSON AND BEAR

;; STREET: 620 NEWPORT CENTER DRIVE SIXTEENTH FLOOR

;; CITY: NEWPORT BEACH

;; STATE: CA

;; COUNTRY: USA

;; ZIP: 92660

;; COMPUTER READABLE FORM:

;; MEDIUM TYPE: Floppy disk

;; COMPUTER: IBM PC compatible

;; OPERATING SYSTEM: PC-DOS/MS-DOS

;; SOFTWARE: Patent in Release #1.0, Version #1.25

;; CURRENT APPLICATION DATA: US/08/100,247

;; FILING DATE: 19930730

;; CLASSIFICATION: 514

;; ATTORNEY/AGENT INFORMATION:

;; NAME: Israel, Ned A.

;; REGISTRATION NUMBER: 29,655

;; REFERENCE/DOCKET NUMBER: O'BRIEN.002A

;; TELEPHONE: 619-235-8550

;; TELEFAX: 619-235-0176

;; INFORMATION FOR SEQ ID NO: 2:

;; SEQUENCE CHARACTERISTICS:

;; LENGTH: 523 amino acids

;; TYPE: amino acid

;; STRANDEDNESS: single

;; TOPOLOGY: linear

;; MOLECULE TYPE: peptide

;; HYPOTHETICAL: NO

;; ANTI-SENSE: NO

;; FRAGMENT TYPE: N-terminal

;; IMMEDIATE SOURCE:

;; CLONE: PROSAPOSIN

US-08-100-247-2

Query Match 100.0%; Score 63; DB 1; Length 523;

Best Local Similarity 100.0%; Pred. No. 0.012;

Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GDVCQDCIQMV 11

Db 194 GDVCQDCIQMV 204

#### RESULT 7

US-08-483-146A-2

;; Sequence 2, Application US/08483146A

;; Patent No. 5696080

;; GENERAL INFORMATION:

;; APPLICANT: O'Brien, John S.

;; APPLICANT: Kishimoto, Yasuo

;; TITLE OF INVENTION: PHARMACEUTICAL COMPOSITIONS

;; TITLE OF INVENTION: COMPRISING PROSAPOSIN AND NEUROTROPHIC PEPTIDES DERIVED

;; TITLE OF INVENTION: THEREFROM

;; NUMBER OF SEQUENCES: 11

;; CORRESPONDENCE ADDRESS:

;; ADDRESSEE: Knobbe, Martens, Olson and Bear

;; STREET: 620 Newport Center Blvd. 16th Floor

;; CITY: Newport Beach

;; STATE: CA

;; COUNTRY: USA

;; ZIP: 92660

;; COMPUTER READABLE FORM:

;; MEDIUM TYPE: Diskette

;; COMPUTER: IBM Compatible

;; OPERATING SYSTEM: DOS

;; SOFTWARE: FastSeq For Windows Version 2.0

;; CURRENT APPLICATION DATA:

;; APPLICATION NUMBER: US/08/483,146A

;; FILING DATE: 07-JUN-1995

;; CLASSIFICATION: 514

;; PRIOR APPLICATION DATA:

;; APPLICATION NUMBER:

;; FILING DATE:

;; ATTORNEY/AGENT INFORMATION:

;; NAME: Israel, Ned A.

;; REGISTRATION NUMBER: 29,655

;; REFERENCE/DOCKET NUMBER: MYELOS.002DV1

;; TELEPHONE: 619-235-8550

;; TELEFAX: 619-235-0176

;; TELEX:

;; INFORMATION FOR SEQ ID NO: 2:

;; SEQUENCE CHARACTERISTICS:

;; LENGTH: 523 amino acids

;; TYPE: amino acid

;; STRANDEDNESS: single

;; TOPOLOGY: linear

;; MOLECULE TYPE: peptide

;; FRAGMENT TYPE: N-terminal

US-08-483-146A-2

Query Match 100.0%; Score 63; DB 1; Length 523;

Best Local Similarity 100.0%; Pred. No. 0.012;

Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GDVCQDCIQMV 11

Db 194 GDVCQDCIQMV 204

#### RESULT 8

US-08-232-513A-3

;; Sequence 3, Application US/08232513A

;; Patent No. 5700909

;; GENERAL INFORMATION:

;; APPLICANT: O'Brien, John S.

;; TITLE OF INVENTION: Prosaoposin and Cytokine-Derived Peptides

;; TITLE OF INVENTION: as Therapeutic Agents

;; NUMBER OF SEQUENCES: 20

;; CORRESPONDENCE ADDRESS:  
;; ADDRESSEE: Campbell & Flores LLP  
;; STREET: 4370 La Jolla Village Drive, Suite 700  
;; CITY: San Diego  
;; STATE: California  
;; COUNTRY: USA  
;; ZIP: 92122  
;; COMPUTER READABLE FORM:  
;; MEDIUM TYPE: Floppy disk  
;; COMPUTER: IBM PC compatible  
;; OPERATING SYSTEM: PC-DOS/MS-DOS  
;; SOFTWARE: Patent in Release #1.0, Version #1.25  
;; CURRENT APPLICATION DATA:  
;; APPLICATION NUMBER: US/08/232,513A  
;; FILING DATE: 21-APR-1994  
;; CLASSIFICATION: 514  
;; PRIOR APPLICATION DATA:  
;; APPLICATION NUMBER: US 08/100,247  
;; FILING DATE: 30-JUL-1993  
;; ATTORNEY/AGENT INFORMATION:  
;; NAME: Campbell, Cathryn A.  
;; REGISTRATION NUMBER: 31,815  
;; REFERENCE/DOCKET NUMBER: P-UD 1643  
;; TELECOMMUNICATION INFORMATION:  
;; TELEPHONE: (619) 535-9001  
;; TELEFAX: (619) 535-8949  
;; INFORMATION FOR SEQ ID NO: 3:  
;; SEQUENCE CHARACTERISTICS:  
;; LENGTH: 523 amino acids  
;; TYPE: amino acid  
;; TOPOLOGY: linear  
;; MOLECULE TYPE: protein  
;; FEATURE:  
;; NAME/KEY: Protein  
;; LOCATION: 1..523  
;; OTHER INFORMATION: /label= Hum\_prosaposin  
US-08-232-513A-3

Query Match 100.0%; Score 63; DB 1; Length 523;  
Best Local Similarity 100.0%; Pred. No. 0.012;  
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GDVQCDCIQMV 11  
|||  
Db 194 GDVQCDCIQMV 204

RESULT 9  
US-08-484-594A-2  
; Sequence 2, Application US/08484594A  
; Patent No. 5714459  
; GENERAL INFORMATION:  
; APPLICANT: O'Brien, John S.  
; APPLICANT: Kishimoto, Yasuo  
; TITLE OF INVENTION: USE OF PROSAPOSIN AND NEUROTROPHIC PEPTIDES  
; TITLE OF INVENTION: DERIVED THEREFROM  
; NUMBER OF SEQUENCES: 11  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Knobbe, Martens, Olson and Bear  
; STREET: 620 Newport Center Drive, Sixteenth Floor  
; CITY: Newport Beach  
; STATE: CA  
; COUNTRY: USA  
; ZIP: 92660  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette  
; COMPUTER: IBM Compatible  
; OPERATING SYSTEM: DOS  
; SOFTWARE: FastSeq for Windows Version 2.0  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/484,594A  
; FILING DATE: 07-JUN-1995  
; CLASSIFICATION: 514

;; PRIOR APPLICATION DATA:  
;; APPLICATION NUMBER: 08/100,247  
;; FILING DATE: 30-JUL-1993  
;; ATTORNEY/AGENT INFORMATION:  
;; NAME: Israelsen, Ned A.  
;; REGISTRATION NUMBER: 29,655  
;; REFERENCE/DOCKET NUMBER: MYELOS.002DV2  
;; TELECOMMUNICATION INFORMATION:  
;; TELEPHONE: 619-235-8550  
;; TELEFAX: 619-235-0176  
;; TELEX:  
;; INFORMATION FOR SEQ ID NO: 2:  
;; SEQUENCE CHARACTERISTICS:  
;; LENGTH: 523 amino acids  
;; TYPE: amino acid  
;; STRANDEDNESS: single  
;; TOPOLOGY: linear  
;; MOLECULE TYPE: peptide  
;; FRAGMENT TYPE: N-terminal  
US-08-484-594A-2

Query Match 100.0%; Score 63; DB 1; Length 523;  
Best Local Similarity 100.0%; Pred. No. 0.012;  
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GDVQCDCIQMV 11  
|||  
Db 194 GDVQCDCIQMV 204

RESULT 10  
US-09-076-258A-2  
; Sequence 2, Application US/09076258A  
; Patent No. 6559124  
; GENERAL INFORMATION:  
; APPLICANT: O'Brien, John S.  
; APPLICANT: Kishimoto, Yasuo  
; TITLE OF INVENTION: PHARMACEUTICAL COMPOSITIONS  
; TITLE OF INVENTION: COMPRISING PROSAPOSIN AND NEUROTROPHIC PEPTIDES DERIVED  
; TITLE OF INVENTION: THEREFROM  
; NUMBER OF SEQUENCES: 11  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Knobbe, Martens, Olson and Bear  
; STREET: 620 Newport Center Blvd. 16th Floor  
; CITY: Newport Beach  
; STATE: CA  
; COUNTRY: USA  
; ZIP: 92660  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette  
; COMPUTER: IBM Compatible  
; OPERATING SYSTEM: DOS  
; SOFTWARE: FastSeq for Windows Version 2.0  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/076,258A  
; FILING DATE:  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/958,970  
; FILING DATE: 28-OCT-97  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/483,146  
; FILING DATE: 07-JUN-1995  
; APPLICATION NUMBER: 08/100,247  
; FILING DATE: 30-JUL-1993  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Israelsen, Ned A.  
; REGISTRATION NUMBER: 29,655  
; REFERENCE/DOCKET NUMBER: MYELOS.2DV1C2  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 619-235-8550  
; TELEFAX: 619-235-0176  
; TELEX:

INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 523 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
FRAGMENT TYPE: N-terminal  
US-09-076-258A-2

Query Match 100.0%; Score 63; DB 4; Length 523;  
Best Local Similarity 100.0%; Pred. No. 0.012;  
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GDVQCDCIQMV 11  
Db 194 GDVQCDCIQMV 204

RESULT 11  
US-08-756-031-2  
Sequence 2, Application US/08756031  
Patent No. 6590074  
GENERAL INFORMATION:  
APPLICANT: O'BRIEN, JOHN S.  
APPLICANT: KISHIMOTO, YASUO  
TITLE OF INVENTION: PROSAPOSIN AS A NEUROTROPHIC FACTOR  
NUMBER OF SEQUENCES: 5  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: KNOBBE, MARTENS, OLSON AND BEAR  
STREET: 620 NEWPORT CENTER DRIVE SIXTEENTH FLOOR  
CITY: NEWPORT BEACH  
STATE: CA  
COUNTRY: USA  
ZIP: 92660  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/756,031  
FILING DATE: 26-NOV-1996  
CLASSIFICATION: 514  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/100,247  
FILING DATE: 30-JUL-1993  
ATTORNEY/AGENT INFORMATION:  
NAME: Israel, Ned A.  
REGISTRATION NUMBER: 29,655  
REFERENCE/DOCKET NUMBER: O'Brien.002A  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 619-235-8550  
TELEFAX: 619-235-0176  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 523 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
HYPOTHETICAL: NO  
ANTI-SENSE: NO  
FRAGMENT TYPE: N-terminal  
IMMEDIATE SOURCE:  
CLONE: PROSAPOSIN  
US-08-756-031-2

Query Match 100.0%; Score 63; DB 4; Length 523;  
Best Local Similarity 100.0%; Pred. No. 0.012;  
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GDVQCDCIQMV 11

Db 194 GDVQCDCIQMV 204

## RESULT 12

US-09-352-548-1  
Sequence 1, Application US/09352548  
Patent No. 6500431  
GENERAL INFORMATION:  
APPLICANT: Gill, Parkash S.  
APPLICANT: Parkash S. Gill, M.D., Inc.  
TITLE OF INVENTION: No. 6500431el Inhibitors of Angiogenesis and Tumor Growth  
FILE REFERENCE: 017986-000410US  
CURRENT APPLICATION NUMBER: US/09/352,548  
CURRENT FILING DATE: 1999-07-12  
EARLIER APPLICATION NUMBER: US 60/092,647  
EARLIER FILING DATE: 1998-07-13  
NUMBER OF SEQ ID NOS: 59  
SOFTWARE: PatentIn Ver. 2.1  
SEQ ID NO 1  
LENGTH: 524  
TYPE: PRT  
ORGANISM: Homo sapiens  
FEATURE: prosaposin  
OTHER INFORMATION: prosaposin  
FEATURE:  
NAME/KEY: PEPTIDE  
LOCATION: (195)..(275)  
OTHER INFORMATION: Saposin B  
US-09-352-548-1

Query Match 100.0%; Score 63; DB 4; Length 524;  
Best Local Similarity 100.0%; Pred. No. 0.012;  
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GDVQCDCIQMV 11  
Db 195 GDVQCDCIQMV 205

## RESULT 13

US-09-352-548-21  
Sequence 21, Application US/09352548  
Patent No. 6500431  
GENERAL INFORMATION:  
APPLICANT: Gill, Parkash S.  
APPLICANT: Parkash S. Gill, M.D., Inc.  
TITLE OF INVENTION: No. 6500431el Inhibitors of Angiogenesis and Tumor Growth  
FILE REFERENCE: 017986-000410US  
CURRENT APPLICATION NUMBER: US/09/352,548  
CURRENT FILING DATE: 1999-07-12  
EARLIER APPLICATION NUMBER: US 60/092,647  
EARLIER FILING DATE: 1998-07-13  
NUMBER OF SEQ ID NOS: 59  
SOFTWARE: PatentIn Ver. 2.1  
SEQ ID NO 21  
LENGTH: 10  
TYPE: PRT  
ORGANISM: Artificial Sequence  
FEATURE:  
OTHER INFORMATION: Description of Artificial Sequence:D2-V11  
OTHER INFORMATION: anti-angiogenic polypeptide  
US-09-352-548-21

Query Match 90.5%; Score 57; DB 4; Length 10;  
Best Local Similarity 100.0%; Pred. No. 0.0027;  
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 DVCQDCIQMV 11  
Db 1 DVCQDCIQMV 10

RESULT 14  
US-09-352-548-48  
; Sequence 48, Application US/09352548  
; Patent No. 6500431  
; GENERAL INFORMATION:  
; APPLICANT: Gill, Parkash S.  
; APPLICANT: Parkash S. Gill, M.D., Inc.  
; TITLE OF INVENTION: No. 6500431el Inhibitors of Angiogenesis and Tumor Growth  
; FILE REFERENCE: 017986-000410US  
; CURRENT APPLICATION NUMBER: US/09/352,548  
; CURRENT FILING DATE: 1999-07-12  
; EARLIER APPLICATION NUMBER: US 60/092,647  
; EARLIER FILING DATE: 1998-07-13  
; NUMBER OF SEQ ID NOS: 59  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 48  
; LENGTH: 16  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence:anti-angiogenic  
; OTHER INFORMATION: polypeptide  
; FEATURE:  
; NAME/KEY: MOD RES  
; LOCATION: (1)-(6)  
; OTHER INFORMATION: Xaa = any amino acid, Xaa at positions 1-6 may be  
; OTHER INFORMATION: present or absent  
US-09-352-548-48

Query Match 90.5%; Score 57; DB 4; Length 16;  
Best Local Similarity 100.0%; Pred. No. 0.0042;  
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 DVCQDCIQMV 11  
Db 7 DVCQDCIQMV 16

RESULT 15  
US-09-352-548-13  
; Sequence 13, Application US/09352548  
; Patent No. 6500431  
; GENERAL INFORMATION:  
; APPLICANT: Gill, Parkash S.  
; APPLICANT: Parkash S. Gill, M.D., Inc.  
; TITLE OF INVENTION: No. 6500431el Inhibitors of Angiogenesis and Tumor Growth  
; FILE REFERENCE: 017986-000410US  
; CURRENT APPLICATION NUMBER: US/09/352,548  
; CURRENT FILING DATE: 1999-07-12  
; EARLIER APPLICATION NUMBER: US 60/092,647  
; EARLIER FILING DATE: 1998-07-13  
; NUMBER OF SEQ ID NOS: 59  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 13  
; LENGTH: 15  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence:anti-angiogenic  
; OTHER INFORMATION: polypeptide  
US-09-352-548-13

Query Match 87.3%; Score 55; DB 4; Length 15;  
Best Local Similarity 90.0%; Pred. No. 0.0082;  
Matches 9; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 GDVCQDCIQM 10  
Db 6 GDVCQDCIQV 15

Search completed: May 5, 2004, 13:16:44  
Job time : 0.480532 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 5, 2004, 13:15:48 ; Search time 1.53561 Seconds  
(without alignments)  
2023.963 Million cell updates/sec

Title: US-09-743-684A-19  
Perfect score: 63  
Sequence: 1 GDVQCDCIQMV 11

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters: 1586107

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : A\_Geneseq\_29Jan04.\*  
1: geneseqp1980s.\*  
2: geneseqp1990s.\*  
3: geneseqp2000s.\*  
4: geneseqp2001s.\*  
5: geneseqp2002s.\*  
6: geneseqp2003as.\*  
7: geneseqp2003bs.\*  
8: geneseqp2004s.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	63	100.0	11	3	AAY58691
2	63	100.0	83	4	AAB31929
3	63	100.0	85	4	AAB31912
4	63	100.0	153	6	ABU70504
5	63	100.0	153	6	ABU70799
6	63	100.0	209	5	ABG70166
7	63	100.0	385	6	ABR39442
8	63	100.0	479	6	ABR39442
9	63	100.0	523	4	ABU31916
10	63	100.0	523	6	ABU05211
11	63	100.0	524	2	AAR70783
12	63	100.0	524	2	AAW85652
13	63	100.0	524	3	AAY58716
14	63	100.0	524	6	ABU79099
15	63	100.0	524	6	ABU05200
16	63	100.0	524	6	ABU05207
17	63	100.0	524	6	ABU05203
18	63	100.0	524	6	ABU07340
19	63	100.0	524	6	ABU05216
20	63	100.0	524	6	ABU05202
21	63	100.0	524	6	ABU05208
22	63	100.0	524	6	ABU05214
23	63	100.0	524	6	ABU05215
24	63	100.0	524	6	ABU05199
25	63	100.0	524	6	ABU05212

ALIGNMENTS

RESULT 1

AAY58691  
ID AAY58691 standard; peptide; 11 AA.

XX AC AAY58691;

XX XX 25-APR-2000 (first entry)

XX DE Antiangiogenic peptide derived from saposin B.

XX KW Antiangiogenic; angiogenesis inhibitor; antitumour; antiproliferative; antimigratory; Kaposi's sarcoma; tumour; human; saposin B; therapy.

XX OS Homo sapiens.

XX FN WO200002902-A1.

XX PD 20-JAN-2000.

XX PF 12-JUL-1999; 99WO-US015772.

XX PR 13-JUL-1998; 98US-0092647P.

XX PA (GILL/) GILL P S.

XX PI Gill PS;

XX DR WPI; 2000-171128/15.

XX PT Saposin B derived peptides, useful as inhibitors of angiogenesis and tumor growth.

XX PS Claim 23; Page 59; 78pp; English.

XX CC The present sequence is that of a claimed peptide, derived from human saposin B, that has antiangiogenic activity. The invention is based on the discovery that saposin B (see AAY58716), previously known to be involved in the hydrolysis of sphingolipids, has potent antiangiogenic and antitumour activity, and also has antiproliferative and antimigratory activity against endothelial cells. This activity is conserved in cryptic polypeptides as small as 5 amino acids (see AAY58684-715), which can be synthetically prepared and used in vitro or in vivo for the treatment of undesired angiogenesis and tumor growth, especially Kaposi's sarcoma (claimed). The polypeptides can also be used in conjunction with cytotoxic moieties to selectively kill certain cell types, e.g. for treatment of cancer, angiofibroma, neovascular glaucoma, arteriovenous malformation, nonunion fracture, arthritis and other connective tissue disorders, Osler-Weber syndrome, atherosclerotic plaque, psoriasis,

Abu05213 Human exp  
Abu05205 Human exp  
Abu05209 Human exp  
Aab31915 Amino aci  
Abp68602 Human pan  
Abu79100 Lip-TAA b  
Abu05206 Human exp  
Abu05204 Human exp  
Abu05210 Human exp  
Aay58693 Antiangio  
Abu70422 Human adi  
Aab58936 Breast an  
Aay58685 Antiangio  
Aay58696 Antiangio  
Aay58704 Antiangio  
Aay58705 Antiangio  
Aay58701 Antiangio  
Aay58697 Antiangio  
Aay58702 Antiangio  
Aaw18588 Universal

CC corneal graft neovascularization, pyogenic granuloma, retrolental  
CC fibroplasia, diabetic retinopathy, scleroderma, haemangioma, trachoma,  
CC vascular adhesions and hypertrophic scars  
XX  
SQ Sequence 11 AA;

Query Match 100.0%; Score 63; DB 3; Length 11;  
Best Local Similarity 100.0%; Pred. No. 0.00095;  
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GDVQCDCIQMW 11  
| | | | | | | | | |  
Db 1 GDVQCDCIQMW 11

RESULT 2  
AAB31929  
ID AAB31929 standard; protein; 83 AA.

AC AAB31929;

DT 15-MAY-2001 (first entry)

DE Amino acid sequence of a human saposin B protein.

Human; perlecan; retinol-binding plasma protein; calgranulin B; vaccine;  
ganglioside GM2 activator; saposin B; degenerative disease; glial cell;  
neurological disease; auto-immune disease; multiple sclerosis; toxicity;  
Alzheimer's disease; Parkinson's disease; amyotrophic lateral sclerosis;  
rheumatoid polyarthrititis; lupus erythematosus; gene therapy.

OS Homo sapiens.

XX WO200105422-A2.

XX 25-JAN-2001.

XX 17-JUL-2000; 2000WO-FR002057.

XX 15-JUL-1999; 99FR-00009372.

XX (INMR ) BIOMERIEUX STELHYS.

XX Roecklin D, Kolbe H, Charles M, Malcus C, Santoro L, Perron H;

XX WPI; 2001-159475/16.

XX Detecting, preventing and treating degenerative, neurological and  
XX autoimmune diseases, particularly multiple sclerosis, using specified  
XX polypeptides or related nucleic acid or ligand.

XX Disclosure; Fig 3; 209pp; French.

XX The present sequence represents a human polypeptide, which is used in the  
XX method of the invention. The specification describes a method which uses  
XX at least one polypeptide or polynucleotide sequence belonging to the  
XX perlecan, precursor of the retinol-binding plasma protein, precursor of  
XX the ganglioside GM2 activator, calgranulin B or saposin B protein  
XX families. The method is used for detecting, preventing or treating a  
XX degenerative, neurological and/or auto-immune disease. The  
XX polynucleotides and polypeptides are used for diagnosis, prognosis,  
XX prevention and treatment of multiple sclerosis (in its various forms and  
XX phases). They may also be useful in cases of e.g. Alzheimer's and  
XX Parkinson's diseases, amyotrophic lateral sclerosis, rheumatoid  
XX polyarthrititis and lupus erythematosus, including use as vaccines and in  
XX gene therapy (expression of sense or antisense sequences). They can also  
XX be used to assess efficacy of potential therapeutic agents, particularly  
XX compounds that reduce or inhibit toxicity towards glial cells

XX Sequence 83 AA;

Query Match 100.0%; Score 63; DB 4; Length 83;  
Best Local Similarity 100.0%; Pred. No. 0.006;

Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 1 GDVQCDCIQMW 11  
| | | | | | | | | |  
Db 1 GDVQCDCIQMW 11

RESULT 3  
AAB31912  
ID AAB31912 standard; protein; 85 AA.

AC AAB31912;

DT 15-MAY-2001 (first entry)

DE Amino acid sequence of a human saposin B protein.

Human; perlecan; retinol-binding plasma protein; calgranulin B; vaccine;  
ganglioside GM2 activator; saposin B; degenerative disease; glial cell;  
neurological disease; auto-immune disease; multiple sclerosis; toxicity;  
Alzheimer's disease; Parkinson's disease; amyotrophic lateral sclerosis;  
rheumatoid polyarthrititis; lupus erythematosus; gene therapy.

OS Homo sapiens.

XX WO200105422-A2.

XX 25-JAN-2001.

XX 17-JUL-2000; 2000WO-FR002057.

XX 15-JUL-1999; 99FR-00009372.

XX (INMR ) BIOMERIEUX STELHYS.

XX Roecklin D, Kolbe H, Charles M, Malcus C, Santoro L, Perron H;

XX WPI; 2001-159475/16.

XX N-PSDB; AAF54720.

XX Detecting, preventing and treating degenerative, neurological and  
XX autoimmune diseases, particularly multiple sclerosis, using specified  
XX polypeptides or related nucleic acid or ligand.

XX Claim 1; Page 169; 209pp; French.

XX The present sequence represents a human polypeptide, which is used in the  
XX method of the invention. The specification describes a method which uses  
XX at least one polypeptide or polynucleotide sequence belonging to the  
XX perlecan, precursor of the retinol-binding plasma protein, precursor of  
XX the ganglioside GM2 activator, calgranulin B or saposin B protein  
XX families. The method is used for detecting, preventing or treating a  
XX degenerative, neurological and/or auto-immune disease. The  
XX polynucleotides and polypeptides are used for diagnosis, prognosis,  
XX prevention and treatment of multiple sclerosis (in its various forms and  
XX phases). They may also be useful in cases of e.g. Alzheimer's and  
XX Parkinson's diseases, amyotrophic lateral sclerosis, rheumatoid  
XX polyarthrititis and lupus erythematosus, including use as vaccines and in  
XX gene therapy (expression of sense or antisense sequences). They can also  
XX be used to assess efficacy of potential therapeutic agents, particularly  
XX compounds that reduce or inhibit toxicity towards glial cells

XX Sequence 85 AA;

Query Match 100.0%; Score 63; DB 4; Length 85;  
Best Local Similarity 100.0%; Pred. No. 0.0061;  
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GDVQCDCIQMW 11  
| | | | | | | | | |  
Db 3 GDVQCDCIQMW 13

```
RESULT 4
ABU70504
ID ABU70504 standard; protein; 153 AA.
XX
XX
AC ABU70504;
XX
XX
DT 10-JUN-2003 (first entry)
XX
XX
DE Human adipocyte Selected Interacting domain, SID, #135.
XX
XX
KW Human; prey; adipocyte; SID; selected interacting domain; anorectic;
KW antidiabetic; protein-protein interaction; diabetes;
KW yeast 2-hybrid assay; metabolic disorder; obesity.
XX
XX
OS Homo sapiens.
XX
XX
PN WO200286122-A2.
XX
XX
PD 31-OCT-2002.
XX
XX
PF 14-MAR-2002; 2002WO-EP003768.
XX
XX
PR 14-MAR-2001; 2001US-0275734P.
XX
XX
PA (HYBR-) HYBRIGENICS.
XX
XX
PI Legrain P, Daviet L;
XX
XX
DR WPI; 2003-103412/09.
DR N-PSDB; ACAS7048.
XX
XX
PT New complex between two interacting proteins in adipocyte cells, useful
PT for identifying selected interacting domains that modulate protein
PT interactions, or for preventing or treating metabolic disorders such as
PT obesity or diabetes.
XX
XX
PS Claim 6; Page 152; 382pp; English.
XX
XX
CC The invention relates to a complex between two interacting proteins in
CC adipocyte cells, given in the specification. The proteins are identified
CC by selecting a bait protein from a known adipocyte marker and then
CC performing a yeast 2-hybrid selection to isolate prey proteins encoded by
CC members of an adipocyte cDNA library. The proteins are designated SID
CC (RTM) (selected interacting domains) proteins. Also included are a
CC polynucleotide encoding a polypeptide in the adipocyte cells, a
CC recombinant host cell expressing at least one of the interacting
CC polypeptides of the complex, selecting a modulating compound in adipocyte
CC cells, a SID (RTM) polypeptide comprising any of the 738 amino acid
CC sequences given in the specification (including its fragment or variant),
CC a SID (RTM) polynucleotide comprising any of the 738 nucleotide sequences
CC given in the specification (including its fragment or variant), a vector
CC comprising the SID (RTM) polynucleotide, a recombinant host cell
CC comprising the vector, a protein chip comprising the polypeptides and a
CC record comprising all or part of the data, listed in the specification.
CC The complex, polypeptides, polynucleotides and compounds are useful for
CC preventing or treating metabolic disorders such as obesity or diabetes.
CC The polynucleotides are useful as probes or primers. The complex is
CC particularly useful for identifying selected interacting domains (SID
CC (RTM)) for screening drugs that modulate the protein interaction, thus
CC exhibiting the therapeutic effect. The present sequence represents a SID
CC (prey) protein of the invention
XX
XX
SQ Sequence 153 AA;
Query Match 100.0%; Score 63; DB 6; Length 153;
Best Local Similarity 100.0%; Pred. No. 0.01;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 GDVQCDCIQMW 11
DB 70 GDVQCDCIQMW 80

RESULT 5
ABU70799
ID ABU70799 standard; protein; 153 AA.
XX
XX
AC ABU70799;
XX
XX
DT 10-JUN-2003 (first entry)
XX
XX
DE Human adipocyte Selected Interacting domain, SID, #430.
XX
XX
KW Human; prey; adipocyte; SID; selected interacting domain; anorectic;
KW antidiabetic; protein-protein interaction; diabetes;
KW yeast 2-hybrid assay; metabolic disorder; obesity.
XX
XX
OS Homo sapiens.
XX
XX
PN WO200286122-A2.
XX
XX
PD 31-OCT-2002.
XX
XX
PF 14-MAR-2002; 2002WO-EP003768.
XX
XX
PR 14-MAR-2001; 2001US-0275734P.
XX
XX
PA (HYBR-) HYBRIGENICS.
XX
XX
PI Legrain P, Daviet L;
XX
XX
DR WPI; 2003-103412/09.
DR N-PSDB; ACAS7343.
XX
XX
PT New complex between two interacting proteins in adipocyte cells, useful
PT for identifying selected interacting domains that modulate protein
PT interactions, or for preventing or treating metabolic disorders such as
PT obesity or diabetes.
XX
XX
PS Claim 6; Page 254; 382pp; English.
XX
XX
CC The invention relates to a complex between two interacting proteins in
CC adipocyte cells, given in the specification. The proteins are identified
CC by selecting a bait protein from a known adipocyte marker and then
CC performing a yeast 2-hybrid selection to isolate prey proteins encoded by
CC members of an adipocyte cDNA library. The proteins are designated SID
CC (RTM) (selected interacting domains) proteins. Also included are a
CC polynucleotide encoding a polypeptide in the adipocyte cells, a
CC recombinant host cell expressing at least one of the interacting
CC polypeptides of the complex, selecting a modulating compound in adipocyte
CC cells, a SID (RTM) polypeptide comprising any of the 738 amino acid
CC sequences given in the specification (including its fragment or variant),
CC a SID (RTM) polynucleotide comprising any of the 738 nucleotide sequences
CC given in the specification (including its fragment or variant), a vector
CC comprising the SID (RTM) polynucleotide, a recombinant host cell
CC comprising the vector, a protein chip comprising the polypeptides and a
CC record comprising all or part of the data, listed in the specification.
CC The complex, polypeptides, polynucleotides and compounds are useful for
CC preventing or treating metabolic disorders such as obesity or diabetes.
CC The polynucleotides are useful as probes or primers. The complex is
CC particularly useful for identifying selected interacting domains (SID
CC (RTM)) for screening drugs that modulate the protein interaction, thus
CC exhibiting the therapeutic effect. The present sequence represents a SID
CC (prey) protein of the invention
XX
XX
SQ Sequence 153 AA;
Query Match 100.0%; Score 63; DB 6; Length 153;
Best Local Similarity 100.0%; Pred. No. 0.01;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 GDVQCDCIQMW 11
DB 74 GDVQCDCIQMW 84
```

RESULT 6  
 ID ABG70166 standard; protein; 209 AA.  
 XX  
 AC ABG70166;  
 XX  
 DT 21-OCT-2002 (first entry)  
 XX  
 DE Human prey protein for Shigella ipaC #30.  
 XX  
 KW Prey protein; ospB; ospD1; ipaD; ipaC; ipaH9.8; ospG; ospC1; Shigella;  
 KW shigellosis; bacillary dysentery; antibacterial; yeast two-hybrid system;  
 KW protein-protein interaction; SID; selected interacting domain; human.  
 XX  
 OS Homo sapiens.  
 XX  
 PN W0200257303-A2.  
 XX  
 PD 25-JUL-2002.  
 XX  
 PF 11-JAN-2002; 2002WO-EP000777.  
 XX  
 PR 12-JAN-2001; 2001US-0261130P.  
 XX  
 PA (HYBR-) HYBRIGENICS.  
 XX  
 PI Legrain P;  
 XX  
 DR WPI; 2002-599706/64.  
 XX  
 DR N-PSDB; ABS51559.  
 XX  
 PT New complex of protein-protein interactions between a bait Shigella  
 PT flexneri polypeptide and a prey mammalian or human placenta polypeptide  
 PT for treating or preventing bacillary dysentery in a mammal or human.  
 XX  
 PS Claim 7; Page 112-113; 162pp; English.  
 XX  
 CC The invention relates to a complex of protein-protein interactions  
 CC between a Shigella flexneri polypeptide (e.g. ospB, ospD1, ipaB, ipaC,  
 CC ipaH9.8, ospG and ospC1) and a mammalian polypeptide defined in the  
 CC specification. The complexes are formed using the yeast two-hybrid  
 CC system. Also included are (1) a recombinant host cell expressing the  
 CC polypeptide defined in the Shigella flexneri polypeptide and a mammalian  
 CC polypeptide defined in the specification; (2) selecting a modulating  
 CC compound that inhibits or activates the protein-protein interactions; (3)  
 CC a modulating compound obtained from the method of (2); (4) a SID  
 CC (selected interacting domain) polypeptide or its fragment or variant  
 CC comprising the human polypeptides appearing as ABG70042-ABG70242; (5) a  
 CC SID polynucleotide or its fragment or variant comprising encoding the  
 CC above polypeptides a vector comprising (5); (6) a recombinant host cell  
 CC containing the vector; and (10) a protein chip comprising Shigella  
 CC flexneri polypeptide and a mammalian polypeptide defined in the  
 CC specification. A pharmaceutical composition comprising the compound,  
 CC polypeptide or polynucleotide is useful for treating or preventing  
 CC shigellosis (bacillary dysentery) in a human or mammal. The present  
 CC sequence represents a human prey protein isolated by the yeast two-hybrid  
 CC assay, forming a complex of the invention with a shigella protein  
 XX  
 SQ Sequence 209 AA;  
 Query Match 100.0%; Score 63; DB 5; Length 209;  
 Best Local Similarity 100.0%; Pred. No. 0.014;  
 Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 Qy 1 GDVQCDCIQWV 11  
 Db 45 GDVQCDCIQWV 55  
 RESULT 7  
 ID ABR41750 standard; protein; 385 AA.  
 XX

AC ABR41750;  
 XX  
 DT 02-JUN-2003 (first entry)  
 XX  
 DE Human DITHP biochemical pathway protein.  
 XX  
 KW Human; dithp; diagnostic and therapeutic polynucleotide; diagnosis;  
 KW cancer; cell proliferative disorder; autoimmune disorder;  
 KW inflammatory disorder; infection; hormonal disorder; metabolic disorder;  
 KW neurological disorder; gastrointestinal disorder; transport disorder;  
 KW connective tissue disorder; drug screening; proteome analysis;  
 KW gene therapy; antisense therapy; genotyping; transgenic animal; knock in;  
 KW disease model; toxicological testing; transcript imaging;  
 KW biochemical pathway.  
 XX  
 OS Homo sapiens.  
 XX  
 PN W0200297031-A2.  
 XX  
 PD 05-DEC-2002.  
 XX  
 PF 27-MAR-2002; 2002WO-US010056.  
 XX  
 PR 28-MAR-2001; 2001US-0279619P.  
 PR 29-MAR-2001; 2001US-0280067P.  
 PR 16-MAY-2001; 2001US-0280068P.  
 PR 17-MAY-2001; 2001US-0291280P.  
 PR 17-MAY-2001; 2001US-0291829P.  
 PR 19-JUN-2001; 2001US-0291849P.  
 PR 19-JUN-2001; 2001US-0299428P.  
 PR 20-JUN-2001; 2001US-0299776P.  
 PR 20-JUN-2001; 2001US-0300001P.  
 XX  
 PA (INCY-) INCYTE GENOMICS INC.  
 XX  
 PI Daffo A, Jones AL, Tran AB, Dahl CR, Gietzen D, Chinn J;  
 PI Duffour GE, Hillman JL, Xu JY, Tuason O, Yap PE, Ameshey SR;  
 PI Daughtery SC, Dam TC, Liu TP, Nguyen DA, Kleefeld Y, Gerstin EH;  
 PI Peralta CH, David WH, Lewis SA, Chen AJ, Panzer SR, Harris B;  
 PI Flores V, Marwaha R, Lo A, Lan RY, Uraahka ME;  
 XX  
 DR WPI; 2003-129518/12.  
 XX  
 DR N-PSDB; ACC46687.  
 XX  
 PT Novel human diagnostic and therapeutic polypeptide useful for identifying  
 PT test compound which specifically binds to a polypeptide encoded by human  
 PT diagnostic and therapeutic polynucleotide, and to induce antibodies.  
 XX  
 PS Claim 27; SEQ ID NO 1285; 591pp; English.  
 XX  
 CC The invention relates to novel human diagnostic and therapeutic  
 CC polynucleotides designated dithp (ACC46080-ACC46749) and to their encoded  
 CC proteins (DITHP; ABR41136-ABR41812). The invention also relates to  
 CC polynucleotide sequences at least 90% identical to the dithp cDNA  
 CC sequences of the invention; recombinant vectors, host cells and  
 CC transgenic organisms comprising a dithp nucleic acid sequence; the  
 CC recombinant production of DITHP proteins; antibodies specific for DITHP  
 CC proteins; microarrays comprising dithp nucleic acid sequences; methods of  
 CC detecting dithp nucleotide and protein sequences; methods of screening  
 CC for compounds which specifically bind a DITHP protein; and methods of  
 CC assessing the toxicity of test compounds using a dithp hybridisation  
 CC probe. Dithp nucleic acid sequences and DITHP proteins may be used in the  
 CC diagnosis of a wide variety of conditions including cancer and other cell  
 CC proliferative disorders; autoimmune or inflammatory disorders; bacterial,  
 CC viral, fungal or parasitic infections; hormonal disorders; metabolic  
 CC disorders; neurological disorders; gastrointestinal disorders; transport  
 CC disorders; and connective tissue disorders. They may also be used to  
 CC screen for modulators of protein activity or gene expression. DITHP  
 CC proteins can additionally be used in analysis of the proteome of a tissue  
 CC or cell type and to induce antibodies. The dithp nucleic acids are  
 CC additionally useful in somatic or germline gene therapy of the disorders  
 CC mentioned above, as a source of antisense sequences, as a source of  
 CC probes and primers, in genotyping and identification of individuals, in

CC the generation of transgenic animal models of human disease or knock in  
CC humanised animals, in toxicological testing, and in transcript imaging.  
CC The present sequence represents a DTHP protein which is involved in a  
CC biochemical pathway. Note: The sequence data for this patent did not form  
CC part of the printed specification, but was obtained in electronic format  
CC directly from WIPO at ftp.wipo.int/pub/published\_pct\_sequences

XX SQ Sequence 385 AA;

Query Match 100.0%; Score 63; DB 6; Length 385;

Best Local Similarity 100.0%; Pred. No. 0.024;

Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GDVCQDCIQMV 11

Db 53 GDVCQDCIQMV 63

RESULT 8

ABR39442

ID ABR39442 standard; protein; 479 AA.

XX AC ABR39442;

XX DT 12-JUN-2003 (first entry)

DE Human GENSET polypeptide clone name SAP-MU-10.

XX GENSET; cytostatic; gene therapy; cancer; transgenic; human.

XX OS Homo sapiens.

XX PN WO2003014151-A2.

XX PD 20-FEB-2003.

XX PF 15-OCT-2001; 2001WO-IB002321.

XX PR 10-AUG-2001; 2001US-0311305P.

XX PR 24-AUG-2001; 2001US-0314734P.

XX PR 07-SEP-2001; 2001US-0318204P.

XX PR 01-OCT-2001; 2001US-0326470P.

XX PA (GSEST ) GENSET SA.

XX PI Bejanin S, Tanaka H;

XX DR WPI; 2003-256539/25.

XX DR N-PSDB; ABZ76265.

XX PT New GENSET gene, useful for preparing a composition for treating GENSET-related disorders.

XX PS Claim 2; Page 288-289; 301pp; English.

XX CC The invention relates to isolated GENSET polynucleotides and encoded polypeptides. The GENSET gene is useful for preparing a composition for treating GENSET-related disorders e.g., cancer. Sequences ABR39423-448 represent the novel GENSET polypeptide sequences

XX SQ Sequence 479 AA;

Query Match 100.0%; Score 63; DB 6; Length 479;

Best Local Similarity 100.0%; Pred. No. 0.03;

Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GDVCQDCIQMV 11

Db 195 GDVCQDCIQMV 205

RESULT 9

AAB31916

ID AAB31916 standard; protein; 523 AA.

XX AC AAB31916;

XX DT 15-MAY-2001 (first entry)

XX DE Amino acid sequence of a human protein.

XX KW Human; perlecan; retinol-binding plasma protein; calgranulin B; vaccine; ganglioside GM2 activator; saposin B; degenerative disease; glial cell; neurological disease; auto-immune disease; multiple sclerosis; toxicity; Alzheimer's disease; Parkinson's disease; amyotrophic lateral sclerosis; rheumatoid polyarthritis; lupus erythematosus; gene therapy.

XX OS Homo sapiens.

XX PN WO200105422-A2.

XX PD 25-JAN-2001.

XX PF 17-JUL-2000; 2000WO-FR002057.

XX PR 15-JUL-1999; 99FR-00009372.

XX PA (INMR ) BIOMERIEUX STELHYS.

XX PI Roeklin D, Kolbe H, Charles M, Malcus C, Santoro L, Perron H;

XX DR WPI; 2001-159475/16.

XX PT Detecting, preventing and treating degenerative, neurological and autoimmune diseases, particularly multiple sclerosis, using specified polypeptides or related nucleic acid or ligand.

XX PS Claim 1; Page 174-175; 209pp; French.

XX CC The present sequence represents a human protein, which is used in the method of the invention. The specification describes a method which uses at least one polypeptide or polynucleotide sequence belonging to the perlecan, precursor of the retinol-binding plasma protein, precursor of the ganglioside GM2 activator, calgranulin B or saposin B protein families. The method is used for detecting, preventing or treating a degenerative, neurological and/or auto-immune disease. The polynucleotides and polypeptides are used for diagnosis, prognosis, prevention and treatment of multiple sclerosis (in its various forms and phases). They may also be useful in cases of e.g. Alzheimer's and Parkinson's diseases, amyotrophic lateral sclerosis, rheumatoid polyarthritis and lupus erythematosus, including use as vaccines and in gene therapy (expression of sense or antisense sequences). They can also be used to assess efficacy of potential therapeutic agents, particularly compounds that reduce or inhibit toxicity towards glial cells

XX SQ Sequence 523 AA;

Query Match 100.0%; Score 63; DB 4; Length 523;

Best Local Similarity 100.0%; Pred. No. 0.032;

Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GDVCQDCIQMV 11

Db 194 GDVCQDCIQMV 204

RESULT 10

ABU05211

ID ABU05211 standard; protein; 523 AA.

XX AC ABU05211;

XX DT 29-JAN-2003 (first entry)

XX DE Human expressed protein tag (EPT) #1877.

XX XX

KW Translational profiling; expressed protein tag; EPT; kinase; phosphatase;  
 KW protease; protease inhibitor; transporter; cytoskeletal protein;  
 KW receptor; transcription factor; cancer; MHC;  
 KW major histocompatibility complex; myeloma; colon cancer; gastric cancer;  
 KW adenocarcinoma; sarcoma; melanoma; lymphoma; leukaemia.

XX Homo sapiens.

OS WO200278524-A2.

PN 10-OCT-2002.

PD 28-MAR-2002; 2002WO-US009671.

PF 28-MAR-2001; 2001US-0279495P.

PR 21-MAY-2001; 2001US-0292544P.

PR 08-AUG-2001; 2001US-0310801P.

PR 01-OCT-2001; 2001US-0326370P.

PR 04-DEC-2001; 2001US-0336780P.

PR 20-FEB-2002; 2002US-0358985P.

XX (ZYCO-) ZYCOS INC.

PA Chicx RM, Tomlinson AJ, Urban RG;

XX WPI; 2003-040607/03.

XX New polypeptides (e.g. kinases, phosphatases, proteases, transporters,  
 PT cytoskeletal proteins, receptors or transcription factors), useful for  
 PT treating cancer, e.g. colon cancer, gastric cancer, sarcoma, lymphoma or  
 PT leukemia.

XX Example 2; SEQ ID NO 1877; 134pp; English.

CC The invention describes a purified polypeptide, which comprises a  
 CC fragment of a kinase, phosphatase, protease, protease inhibitor,  
 CC transporter, cytoskeletal protein, receptor or transcription factor. The  
 CC polypeptide is useful as an immunogenic composition for eliciting in a  
 CC mammal an immunogenic response directed against any of the purified  
 CC polypeptides. The purified polypeptide, or the antibody that binds to this  
 CC polypeptide, is useful for treating cancer. The polypeptide is also  
 CC useful for identifying compounds that binds to a naturally processed  
 CC class I or class II MHC-binding polypeptide. The polypeptides and  
 CC polynucleotides are particularly useful for treating or preventing  
 CC myeloma, colon cancer, gastric cancer, adenocarcinoma, sarcoma, melanoma,  
 CC lymphoma or leukaemia. These are also useful for screening agents for  
 CC treating the above mentioned diseases. This sequence represents an  
 CC expressed protein tag (EPT) isolated from human tissue for translational  
 CC profiling. Note: This sequence does not appear in the printed  
 CC specification but was obtained in electronic format directly from WIPO at  
 CC ftp.wipo.int/pub/published\_pct\_sequences

XX Sequence 523 AA;

Query Match 100.0%; Score 63; DB 6; Length 523;  
 Best Local Similarity 100.0%; Pred. No. 0.032;  
 Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 GDVQCDCIQMV 11  
 |||||

Db 194 GDVQCDCIQMV 204  
 |||||

RESULT 11

ID AAR70783 standard; protein; 524 AA.

XX AAR70783;

XX 25-MAR-2003 (revised)

DT 30-AUG-1995 (first entry)

XX Prosaposin.

DE

XX

KW

XX

KW

XX

KW

XX

OS

XX

PN

XX

PD

XX

PF

XX

PR

XX

PR

XX

PA

XX

PI

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX Saposin-C; neuron; myelination; nervous system; neuroblastoma;  
 KW neurotrophic peptide; multiple sclerosis; leukoencephalitis;  
 KW adrenal leukodystrophy; prosaposin.

OS Homo sapiens.

PN WO9503821-A1.

PD 09-FEB-1995.

PF 28-JUL-1994; 94WO-US008453.

PR 30-JUL-1993; 93US-00100247.

PR 21-APR-1994; 94US-00232513.

PA (OBRI/) O'BRIEN J S.

PI O'brien JS, Kishimoto Y;

XX WPI; 1995-082029/11.

XX N-PSDB; AAQ85355.

XX Stimulating neural cell out-growth and myelination - with pro:saposin,  
 PT saposin C or new neurotrophic peptide(s) from cytokine(s), for treating  
 PT nervous system diseases.

XX Disclosure; Page 30-32; 50pp; English.

CC The peptide given in AAR70773, corresponding to amino acids 8-29 of human  
 CC saposin-C (AAR70784), promotes neurite outgrowth in vitro. A consensus  
 CC sequence was determined by comparing the peptide with hematopoietic and  
 CC neurotrophic cytokines, and neurotrophic peptides (AAR70774-82) were  
 CC identified in the AB loop of human ciliary neurotrophic factor,  
 CC interleukins-6, -2, -3 and -gamma, erythropoietin and leukocyte  
 CC inhibitory factor, and in helix C of human interleukin-1-beta and  
 CC oncostatin-M. Prosaposin (AAR70783) and saposin-C also promoted nerve  
 CC cell myelination ex vivo. (Updated on 25-MAR-2003 to correct PN field.)  
 CC (Updated on 25-MAR-2003 to correct PA field.) (Updated on 25-MAR-2003 to  
 CC correct PI field.)

XX Sequence 524 AA;

Query Match 100.0%; Score 63; DB 2; Length 524;  
 Best Local Similarity 100.0%; Pred. No. 0.032;  
 Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 GDVQCDCIQMV 11  
 |||||

Db 195 GDVQCDCIQMV 205  
 |||||

RESULT 12

AAR85652

ID AAR85652 standard; protein; 524 AA.

XX AAR85652;

XX 19-JUL-1999 (first entry)

XX Human prosaposin N-terminal peptide.

XX Prosaposin; saposin; prosapides; prosaposin receptor agonists; PRA;  
 KW peripheral nervous system; central nervous system; PNS; CNS; Akt; Bcl-2;  
 KW therapy; treatment; apoptosis; caspase; tumour necrosis factor; TNF;  
 KW cytokine; interferon gamma; IFN; inflammation; rheumatoid arthritis;  
 KW Crohn's disease; irritable bowel syndrome; asthma; cardiac infarction;  
 KW congestive heart failure; multiple sclerosis;  
 KW acute disseminated inflammatory leukoencephalitis;  
 KW progressive multifocal leukoencephalitis; Alzheimer's disease;  
 KW Parkinson's disease; amyotrophic lateral sclerosis; Huntington's disease;  
 KW ischemic heart disease; Guillain-Barre disease; alopecia; AIDS dementia;  
 KW cerebral malaria; HTLV; neuropathy;

inflammatory neurodegenerative disease; toxin-induced liver disease.

XX KW Homo sapiens.  
XX OS  
XX PN W09912559-A1.  
XX XX 18-MAR-1999.  
XX XX 09-SEP-1998; 98WO-US019216.  
XX XX 09-SEP-1997; 97US-0058352P.  
XX PR 04-JUN-1998; 98US-0088129P.  
XX XX (REGC ) UNIV CALIFORNIA.  
XX PI O'brien JS;  
XX DR WPI; 1999-229139/19.  
XX DR N-PSDB; AAX08488.  
XX PT Use of prosaposin receptor agonist.  
XX PS Claim 7; Fig 2; 90pp; English.  
XX XX Prosaposin is a 70kDa glycoprotein which is proteolytically processed to  
CC generate saposins A, B, C and D, all of which are similar to each other  
CC and have a similar placement of six cysteines, a glycosylation site and  
CC conserved proline residues. Prosaposin, saposin C and prosaposin derived  
CC peptides (prosaptides) have therapeutic applications in promoting  
CC recovery after toxic, traumatic, myocardial ischaemic, degenerative and  
CC inherited lesions to the peripheral and central nervous system.  
CC Prosaposin receptor agonists (PRAs) inhibit proinflammatory cytokine-  
CC induced apoptosis by activation of the Ser/Thr protein kinase Akt. Akt  
CC dissociates complexes of Bcl-2 family members, such as BAD-Bcl-2,  
CC releasing Bcl-2 and its family members which inhibit caspases, thereby  
CC inhibiting apoptosis. An additional mechanism whereby PRAs inhibit  
CC apoptosis is by blocking activation of JNK, a proapoptotic signaling  
CC component. Within several minutes after binding to the receptor, PRAs  
CC block JNK activation induced by tumor necrosis factor-alpha (TNF alpha).  
CC The activation of JNK by TNF alpha is another well known mechanism for  
CC TNF alpha-induced, as well as other proinflammatory cytokine-induced  
CC apoptosis. The method can be used for inhibiting apoptosis which is  
CC caspase-mediated or induced by a proinflammatory cytokine, for example  
CC TNF alpha or interferon-gamma. It can be used for inhibiting apoptosis  
CC associated with a disorder such as e.g. rheumatoid arthritis, Crohn's  
CC disease, irritable bowel syndrome, asthma, cardiac infarction, congestive  
CC heart failure, multiple sclerosis, acute disseminated inflammatory  
CC leukoencephalitis, progressive multifocal leukoencephalitis, Alzheimer's  
CC disease, Parkinson's disease, amyotrophic lateral sclerosis, Huntington's  
CC disease, ischemic heart disease, Guillain-Barre disease, traumatic brain  
CC injury, traumatic spinal cord injury, alopecia, AIDS dementia, cerebral  
CC malaria, HTLV, neuropathy, inflammatory neurodegenerative disease, and  
CC toxin-induced liver disease. This 524 N-terminal peptide of prosaposin  
CC also acts as a prosaposin receptor agonist  
XX SQ Sequence 524 AA;  
Query Match 100.0%; Score 63; DB 2; Length 524;  
Best Local Similarity 100.0%; Pred. No. 0.032;  
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Oy 1 GDVCQDCIQMV 11  
Db 195 GDVCQDCIQMV 205  
RESULT 13  
AAY58716  
ID AAY58716 standard; protein; 524 AA.  
XX XX  
AC AAY58716;  
XX DT 25-APR-2000 (first entry)

XX Human prosaposin.  
XX DE  
XX KW Prosaposin; saponin B; antiangiogenic; angiogenesis inhibitor;  
KW antitumour; antiproliferative; antimigratory; Kaposi's sarcoma; tumour;  
KW human; therapy.  
XX OS Homo sapiens.  
XX PH Key Location/Qualifiers  
XX XX 195..275  
XX FT Protein /note= "mature saposin B"  
XX FT Peptide 195..205  
XX FT Claim 23" /note= "specifically claimed antiangiogenic peptide of  
XX FT Peptide 196..200 /note= "specifically claimed antiangiogenic peptide of  
XX FT Claim 4"  
XX PN W0200002902-A1.  
XX XX 20-JAN-2000.  
XX PD  
XX PF 12-JUL-1999; 99WO-US015772.  
XX XX 13-JUL-1998; 98US-0092647P.  
XX PR (GILL/) GILL P S.  
XX PA Gill PS;  
XX PI  
XX DR WPI; 2000-171128/15.  
XX PT Saposin B derived peptides, useful as inhibitors of angiogenesis and  
XX tumor growth.  
XX PS Disclosure; Page 18; 78pp; English.  
XX XX The present sequence is that of human prosaposin, a precursor of saposin  
CC B. The invention is based on the discovery that saposin B, previously  
CC known to be involved in the hydrolysis of sphingolipids, has potent  
CC antiangiogenic and antitumour activity, and also has antiproliferative  
CC and antimigratory activity against endothelial cells. This activity is  
CC conserved in cryptic polypeptides as small as 5 amino acids (see AAY58684  
CC -715), which can be synthetically prepared and used in vitro or in vivo  
CC for the treatment of undesired angiogenesis and tumor growth, especially  
CC Kaposi's sarcoma (claimed). The polypeptides can also be used in  
CC conjunction with cytotoxic moieties to selectively kill certain cell  
CC types, e.g. for treatment of cancer, angiofibroma, neovascular glaucoma,  
CC arteriovenous malformation, nonunion fracture, arthritis and other  
CC connective tissue disorders, Osler-Weber syndrome, atherosclerotic  
CC plaque, psoriasis, corneal graft neovascularization, pyogenic granuloma,  
CC retrolental fibroplasia, diabetic retinopathy, scleroderma, haemangioma,  
CC trachoma, vascular adhesions and hypertrophic scars  
XX SQ Sequence 524 AA;  
Query Match 100.0%; Score 63; DB 3; Length 524;  
Best Local Similarity 100.0%; Pred. No. 0.032;  
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Oy 1 GDVCQDCIQMV 11  
Db 195 GDVCQDCIQMV 205  
RESULT 14  
ABU79099  
ID ABU79099 standard; protein; 524 AA.  
XX XX  
AC ABU79099;  
XX DT 18-JUN-2003 (first entry)

XX DE Lip-TAA binding protein, Proasoposin.  
XX KW Superantigen; SAg; staphylococcal enterotoxin; tumour; cancer; apoptosis;  
XX KW gene therapy; mammalian cell receptor; cytosolic;  
XX KW tumour associated lipid; energy; T cell; antigen presenting cell; APC;  
XX KW tumoricidal immunocyte; antitumour.  
XX OS Unidentified.  
XX FN US2002177551-A1.  
XX XX 28-NOV-2002.  
XX PD 30-MAY-2001; 2001US-00870759.  
XX PF 31-MAY-2000; 2000US-0208128P.  
XX PR (TERM/) Terman D S.  
XX PA Terman DS;  
XX PI WPI; 2003-361759/34.  
XX DR A mammalian cell receptor, useful in the treatment of cancer by binding  
XX PT to tumor associated lipids where the binding induces energy or apoptosis  
XX PT in T cells and antigen presenting cells.  
XX PS Disclosure; Page; 167pp; English.  
XX XX  
XX CC The invention relates to a mammalian cell receptor, useful in the  
XX CC treatment of cancer, which binds to tumour associated lipids and induces  
XX CC energy or apoptosis in the T cells and antigen presenting cells (APCs).  
XX CC Also included are a mammalian cell useful in the treatment of cancer  
XX CC where the receptor which binds tumour associated lipids and induces  
XX CC cellular inactivation or death is deleted or functionally deactivated,  
XX CC producing (M1) a tumoricidal immunocyte population in vivo in a mammal  
XX CC (by allowing tumour associated lipids to contact immunocytes in which  
XX CC receptors for immunosuppressive fatty acids, ceramides, glycolipids,  
XX CC sphingolipids, glycosphingolipids, phosphosphingolipids, gangliosides,  
XX CC sialylated glycans, lipopeptides and proteoglycolipids are inactivated or  
XX CC deleted), a construct useful in the treatment of cancer comprising a  
XX CC superantigen (SAg) nucleotide inserted into a virus, a mammalian T cell  
XX CC useful in the treatment of cancer (where an adaptor protein which  
XX CC inhibits T cell activation by tumour associated antigens is deleted or  
XX CC functionally deactivated), a composition useful in the treatment of  
XX CC cancer (comprising a lipid raft conjugated to a superantigen), producing  
XX CC allowing tumour associated lipids to contact immunocytes, in which  
XX CC tumoricidal immunocyte population, and administering the tumoricidally  
XX CC activated immunocytes to the host), producing (M3) a tumoricidal APC  
XX CC population ex vivo in a mammal (by allowing a tumour associated lipid to  
XX CC contact APCs in which receptors for the tumour associated lipids are  
XX CC inactivated or deleted to produce a tumoricidally activated population,  
XX CC and administering APCs to the host), producing a tumoricidal T cell  
XX CC population ex vivo in a mammal (by allowing a tumour associated lipids to  
XX CC contact T cells, in which adaptor proteins, which inhibit T cell  
XX CC activation by tumour associated antigens, are deleted or functionally  
XX CC deactivated to produce a tumoricidally activated population of T cells, and  
XX CC administering the tumoricidally activated T cells to the host, or  
XX CC allowing a superantigen-lipid raft to contact T cells ex vivo, and  
XX CC administering the tumoricidally activated T cells to the host), treating  
XX CC (M5) cancer in a mammal (by administering a lipid binding molecule which  
XX CC binds immunosuppressive tumour associated lipids in vivo), producing (M6)  
XX CC a tumoricidal T cell population in vivo in a mammal (by allowing a  
XX CC tumour associated antigen to contact immunocytes in which adaptor  
XX CC proteins which inhibit T cell activation by tumour associated antigens  
XX CC are deleted or functionally deactivated) and producing (M7) a  
XX CC tumoricidal T cell population ex vivo in a mammal comprising allowing a  
XX CC superantigen-lipid raft conjugate to contact immunocytes in vivo. The  
XX CC receptors, methods and compositions are useful for treating cancers and  
XX CC tumours. Bacterial superantigens are co-administered or administered as

CC fusion constructs with anti-tumour proteins or motifs. The present  
CC sequence represents a tumour antigen or a motif identifying a tumour  
CC antigen, which can be functionally deactivated in the method of the  
CC invention. Note: The sequence data for this patent did not form part of  
CC the printed specification, but was obtained in electronic format from the  
CC US patent office website at  
CC "seqdata.uspto.gov/sequence.html?DocID=20020177551"  
XX SQ Sequence 524 AA;  
XX  
XX Query Match 100.0%; Score 63; DB 6; Length 524;  
XX Best Local Similarity 100.0%; Pred. No. 0.032;  
XX Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
XX  
XX Qy 1 GDVCQDCIQMW 11  
XX Db 195 GDVCQDCIQMW 205  
XX  
XX RESULT 15  
XX ABU05200  
XX ID ABU05200 standard; protein; 524 AA.  
XX AC ABU05200;  
XX XX  
XX DT 29-JAN-2003 (first entry)  
XX XX  
XX DE Human expressed protein tag (EPT) #1866.  
XX KW  
XX KW Translational profiling; expressed protein tag; EPT; kinase; phosphatase;  
XX KW protease; protease inhibitor; transporter; cytoskeletal protein;  
XX KW receptor; transcription factor; cancer; MHC;  
XX KW major histocompatibility complex; myeloma; colon cancer; gastric cancer;  
XX KW adenocarcinoma; sarcoma; melanoma; lymphoma; leukaemia.  
XX OS Homo sapiens.  
XX OS  
XX PN WO200278524-A2.  
XX XX  
XX PD 10-OCT-2002.  
XX XX  
XX PF 28-MAR-2002; 2002WO-US009671.  
XX PR 28-MAR-2001; 2001US-0279495P.  
XX PR 21-MAY-2001; 2001US-0292544P.  
XX PR 08-AUG-2001; 2001US-0310801P.  
XX PR 01-OCT-2001; 2001US-0326370P.  
XX PR 04-DEC-2001; 2001US-0336780P.  
XX PR 20-FEB-2002; 2002US-0358985P.  
XX PA (ZYCO-) ZYCO INC.  
XX XX  
XX PI Chicx RM, Tomlinson AJ, Urban RG;  
XX XX  
XX DR WPI; 2003-040607/03.  
XX XX  
XX PT New polypeptides (e.g. kinases, phosphatases, proteases, transporters,  
XX PT cytoskeletal proteins, receptors or transcription factors), useful for  
XX PT treating cancer, e.g. colon cancer, gastric cancer, sarcoma, lymphoma or  
XX XX leukemia.  
XX PS Example 2; SEQ ID NO 1866; 134pp; English.  
XX XX  
XX CC The invention describes a purified polypeptide, which comprises a  
XX CC fragment of a kinase, phosphatase, protease, protease inhibitor,  
XX CC transporter, cytoskeletal protein, receptor or transcription factor. The  
XX CC polypeptide is useful as an immunogenic composition for eliciting in a  
XX CC mammal an immunogenic response directed against any of the purified  
XX CC polypeptide. The purified polypeptide, or the antibody that binds to this  
XX CC polypeptide, is useful for treating cancer. The polypeptide is also  
XX CC useful for identifying compounds that binds to a naturally processed  
XX CC class I or class II MHC-binding polypeptide. The polypeptides and  
XX CC polynucleotides are particularly useful for treating or preventing



CC myeloma, colon cancer, gastric cancer, adenocarcinoma, sarcoma, melanoma,  
CC lymphoma or leukaemia. These are also useful for screening agents for  
CC treating the above mentioned diseases. This sequence represents an  
CC expressed protein tag (EPT) isolated from human tissue for translational  
CC profiling. Note: This sequence does not appear in the printed  
CC specification but was obtained in electronic format directly from WIPO at  
CC ftp.wipo.int/pub/published\_pct\_sequences

XX

SQ Sequence 524 AA;

Query Match 100.0%; Score 63; DB 6; Length 524;  
Best Local Similarity 100.0%; Pred. No. 0.032;  
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GDVCQDCIQMV 11

|||||

Db 195 GDVCQDCIQMV 205

Search completed: May 5, 2004, 13:31:04  
Job time : 1.53561 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 5, 2004, 13:25:43 ; Search time 52.2507 Seconds  
(without alignments)  
3164.197 Million cell updates/sec

Title: US-09-743-684A-1  
Perfect score: 2789  
Sequence: 1 MYALFLIASLLGALAGPL.....NTETAAQCNVAECHKRHVWN 524

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 1017041 seqs, 315518202 residues

Total number of hits satisfying chosen parameters: 1017041

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

- Database : SPTREMBL 25:\*
- 1: sp\_archaea:\*
  - 2: sp\_bacteria:\*
  - 3: sp\_fungi:\*
  - 4: sp\_human:\*
  - 5: sp\_invertebrate:\*
  - 6: sp\_mammal:\*
  - 7: sp\_mhc:\*
  - 8: sp\_organelle:\*
  - 9: sp\_phage:\*
  - 10: sp\_plant:\*
  - 11: sp\_rodent:\*
  - 12: sp\_virus:\*
  - 13: sp\_vertebrate:\*
  - 14: sp\_unclassified:\*
  - 15: sp\_rvirus:\*
  - 16: sp\_bacteriap:\*
  - 17: sp\_archaeap:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1950	69.9	554	11 Q8BFQ1	Q8bfq1 mus musculus
2	1442	51.7	512	13 Q7SY70	Q7sy70 xenopus lae
3	1414	50.7	520	13 Q8UVZ4	Q8uvz4 brachydanio
4	1343	48.2	522	13 Q9DG82	Q9dg82 brachydanio
5	1099.5	39.4	525	11 Q8C1C1	Q8c1c1 mus musculus
6	786	28.2	449	11 Q8BJV5	Q8bjv5 mus musculus
7	772	27.7	402	11 Q8C1N0	Q8c1n0 mus musculus
8	552	19.8	953	5 Q9Y125	Q9y125 drosophila
9	545	19.5	241	4 Q9N714	Q9n7t4 homo sapien
10	517	18.5	121	6 P79254	P79254 ovis aries
11	511	18.3	965	5 Q15997	Q15997 bombyx mori
12	432.5	15.5	876	5 Q8IMH4	Q8imh4 drosophila
13	349.5	12.5	378	11 Q35489	Q35489 cavia porce
14	343.5	12.3	458	5 Q95X02	Q95x02 naegleria f
15	343.5	12.3	484	5 Q9BKMI	Q9bkmi naegleria f
16	334.5	12.0	374	6 Q9TU81	Q9tu81 ovis aries

17	327.5	11.7	441	5 Q9U9P3	Q9u9p3 drosophila
18	292.5	10.5	370	6 P79333	P79333 oryctolagus
19	288	10.3	307	5 Q9BKM2	Q9bkm2 naegleria f
20	279	10.0	294	5 Q95X03	Q95x03 naegleria f
21	279	10.0	456	5 Q94472	Q94472 dictyosteli
22	251	9.0	316	5 Q86PA4	Q86pa4 drosophila
23	220.5	7.9	429	5 Q18276	Q18276 caenorhabdi
24	213	7.6	213	10 Q9SCT5	Q9sct5 arabidopsis
25	203.5	7.3	217	10 Q9LZW6	Q9lzw6 arabidopsis
26	187.5	6.7	243	6 Q9TT05	Q9tt05 ovis aries
27	180	6.5	200	5 Q86JD8	Q86jd8 dictyosteli
28	179.5	6.4	240	10 Q9AS89	Q9as89 oryza sativ
29	176.5	6.3	402	5 Q18279	Q18279 caenorhabdi
30	176	6.3	228	6 Q9BDZ9	Q9bdz9 ovis aries
31	173.5	6.2	174	10 Q9M614	Q9m614 vitis ripar
32	173.5	6.2	507	10 Q9FRW7	Q9frw7 nepenthes a
33	172	6.2	188	5 Q86KA8	Q86ka8 dictyosteli
34	171.5	6.1	486	10 Q38934	Q38934 arabidopsis
35	171.5	6.1	506	10 Q6S390	Q6s390 arabidopsis
36	165.5	5.9	506	10 Q39311	Q39311 brassica na
37	163.5	5.9	273	10 Q9LUX5	Q9lux5 pyrus pyrif
38	163.5	5.9	513	10 Q8VYL3	Q8vyl3 arabidopsis
39	162.5	5.8	513	10 Q41713	Q41713 vigna ungui
40	162	5.8	514	10 Q8L6A9	Q8l6a9 theobroma c
41	161	5.8	107	10 Q8HOR9	Q8hor9 cynara card
42	157	5.6	512	10 Q04593	Q04593 arabidopsis
43	156.5	5.6	514	10 Q94IA2	Q94ia2 glycine max
44	155.5	5.6	205	6 Q9N275	Q9n275 ovis aries
45	155.5	5.6	292	10 Q43407	Q43407 brassica ol

ALIGNMENTS

RESULT 1

ID	Q8BFQ1	PRELIMINARY;	PRT;	554 AA.
AC	Q8BFQ1;			
DT	01-MAR-2003 (T-EMBLrel. 23, Created)			
DT	01-MAR-2003 (T-EMBLrel. 23, Last sequence update)			
DT	01-OCT-2003 (T-EMBLrel. 25, Last annotation update)			
DE	Prosaposin.			
GN	PSAP.			
OS	Mus musculus (Mouse).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Musinae; Mus.			
OX	NCBI_TaxID=10090;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN=NOD; TISSUE=Kidney, and Thymus;			
RX	MEDLINE=22354683; PubMed=12466851;			
RA	The FANTOM Consortium.			
RA	the RIKEN Genome Exploration Research Group Phase I & II Team;			
RT	"Analysis of the mouse transcriptome based on functional annotation of			
RT	60,770 full-length cDNAs."			
RL	Nature 420:563-573(2002).			
DR	EMBL; AK088369; BAC40308.1; --			
DR	MGD; MGI:97783; Psap.			
DR	GO; GO:0005764; C:lysosome; IEA.			
DR	GO; GO:0006655; P:sphingolipid metabolism; IEA.			
DR	InterPro; IPR003119; Sapa.			
DR	InterPro; IPR007856; SapaB_1.			
DR	InterPro; IPR008138; SapaB_2.			
DR	InterPro; IPR008140; SapaB_sub.			
DR	InterPro; IPR008373; Saposin.			
DR	Pfam; PF02199; SAPA; 2.			
DR	Pfam; PF05184; SapaB_1; 4.			
DR	Pfam; PF03489; SapaB_2; 4.			
DR	PRINTS; PR01797; SAPOSIN.			
DR	ProDom; PD001732; SapaB_sub; 3.			
DR	SMART; SM00162; SAPA; 2.			

DR SMART; SM00118; SAPB; 4.  
 SQ SEQUENCE 554 AA; 61050 MW; PF58DB79C7C0C018 CRC64;

Query Match 69.9%; Score 1950; DB 11; Length 554;  
 Best Local Similarity 64.0%; Pred. No. 4.4e-142;  
 Matches 355; Conservative 78; Mismatches 90; Indels 32; Gaps 2;

Qy 1 MYALFLLASLLGALAGPVLGLKECTGSAVQCNVKTASDCGAVKHCLQTVNKPVTKS 60  
 Db 1 MYALALFASILLATATSTFVDPKTCGGSVAIVCRDVKTAVDCGAVKHCOQMWWSKPTAKS 60

Qy 61 LPDICKDVVTAAGDMLKONATEBEILVLEKTCDWLPKPNMSASCKETVDSYLPVILDI 120  
 Db 61 LPDICKVTVTEAGNLKONATQEEILHYLEKTECIWHDSLSASCKEVDYSLPVILDM 120

Qy 121 IKGEMSRPGVCSALNLCESLQKHLAEHLNQHOKLESNKIPELDMTEVWAPFMANIPLLY 180  
 Db 121 IKGEMNPGVCSALNLCESLQKHLAEHLNQHOKLESNKIPEDMARVWAPFMANIPLLY 179

Qy 181 PQDPRSKPQKNGDVCQDCIQMVTDTIQTAVRTNSTFVQALVEHVKECDRLGPGMADI 240  
 Db 180 PQDPRSQPKANEDVCQDKMLVSDVQTAVKTNSSTFIOGFVDHVKECDRLGPGVSDI 239

Qy 241 CKNYISOYSEIATQMMHMQPKIEICALVGFCEVEMPMQTLVPKAVSKNVPALBELVE 300  
 Db 240 CKNYDQYSEVVCQMLMHMQPKIEICVLAGFCNEVRVPKTLVPATETIKNLPALLENMD 299

Qy 301 PIKKEVPAKSDVYCEVCEFLVKEVTKLIDNNKTEKEILDADFCKWCKLPKSLSEBCEQV 360  
 Db 300 PVEQNVLQAHNVILQTCQFVNNKSEILVNNATEELLVGLSNACALLPDPARTKQEV 359

Qy 361 VDTYGSLSILLESVPELVCSMLHLCSG----- 390  
 Db 360 VGTGFSLLDIFIHEVNPSSLCGVIGLCAARPELVEALEQAPAIVSALLKEPTPKQPA 419

Qy 391 -TRLPALTIVHTVPKDGDFCECKLVGYLDRLNLEKSTKQEIILAELEKGSFLPDPYQK 449  
 Db 420 QPKQALPAHPVPQKNGGFCVCKLVLYLEHNLKSTKEEILAELEKGSFLPDPYQK 479

Qy 450 QCDQFAEYEPVLBIELVMDPFCVCLIKIGACPSAKHKLIGTEKCIWGPSYQCONTEA 509  
 Db 480 QCDQFAEYEPVLBIELVMDPFCVCSKIGVCSAYKLLJLGTCKVGFSGYQCONMETH 539

Qy 510 AQCNVAHECHKRHVN 524  
 Db 540 ARCNVAHECHKRHVN 554

RESULT 2  
 ID Q7SY70 PRELIMINARY; PRT; 512 AA.  
 AC Q7SY70;  
 DT 01-OCT-2003 (TrEMBLrel. 25, Created)  
 DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)  
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
 DE Hypothetical protein.  
 OS Xenopus laevis (African clawed frog).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae; Pipidae;  
 OC Xenopodinae; Xenopus.  
 OX NCBI\_TaxID=8355;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Whole;  
 RX MEDLINE=22341132; PubMed=12454917;  
 RA Klein S.L., Strausberg R.L., Wagner L., Pontius J., Clifton S.W.,  
 RA Richardson P.;  
 RT "Genetic and genomic tools for Xenopus research: The NIH Xenopus  
 RT initiative.";  
 RL Dev. Dyn. 225:384-391 (2002).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Whole;

RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,  
 RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,  
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,  
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Haie H.,  
 RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,  
 RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,  
 RA Brownstein M.J., Udwin T.B., Toshiyuki S., Carninci P., Prange C.,  
 RA Raha S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,  
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,  
 RA Villalón D.K., Muzny K.C., Sodergren E.J., Lu X., Gibbs R.A.,  
 RA Fahey J., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A.,  
 RA Whiting M., Madan A., Young A.C., Shevchenko V., Bouffard G.G.,  
 RA Blakeley R.W., Touchman J.W., Green E.D., Dickson M.C.,  
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,  
 RA Krzywinski M.I., Skalius D., Smailus D.E., Schnerch A., Schein J.E.,  
 RA Jones S.J., Marra M.A.;  
 RT "Generation and initial analysis of more than 15,000 full-length human  
 RT and mouse cDNA sequences.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Whole;  
 RC Klein S., Strausberg R.;  
 RL Submitted (JUL-2003) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; BC054988; AAH54988.1; --  
 KW Hypothetical protein.  
 SQ SEQUENCE 512 AA; 57555 MW; 57CFASE2093F6FB1 CRC64;

Query Match 51.7%; Score 1442; DB 13; Length 512;  
 Best Local Similarity 46.3%; Pred. No. 6.4e-103;  
 Matches 243; Conservative 131; Mismatches 137; Indels 14; Gaps 7;

Qy 1 MYALFLLASLLGALAGPVLGLKECTGSAVQCNVKTASDCGAVKHCLQTVNKPVTKS 60  
 Db 1 MKQLAVLFCALALVAATFLFTEQCAKGFVWCNVRTASQCGAVKHCOQMWWSKPTAKS 60

Qy 61 LPDICKDVVTAAGDMLKONATEBEILVLEKTCDWLPKPNMSASCKEIVDSYLPVILDI 120  
 Db 61 LPDCKVKEIITVLGNFMKDNITQGEIKDYLNKVCDFPDPGLAATCKQEVSDYFISVLN 120

Qy 121 IKGEMSRPGVCSALNLCESLQKHLAEHLNQHOKLESNKIPELDMTEVWAPFMANIPLLY 180  
 Db 121 LQQLSNPGLVLCSSGLGCTSLQRHLASLKQPKQLLNNEIPDVSASKLYPFIIVNPVQLY 180

Qy 181 PQDPRSKPQKNGDVCQDCIQMVTDTIQTAVRTNSTFVQALVEHVKECDRLGPGMADI 240  
 Db 181 PQD--KTPKEPK--TGDI--CNDCTTLLISDVQALRSNSSFSSKLVDFHMQECLNLDPTWAE 237

Qy 241 CKNYISOYSEIATQMMHMQPKIEICALVGFCEVEMPMQTLVPKAVSKNVPALBELVE 300  
 Db 238 CKSYNQVADIATQVLLQMQPKLQCGNAGFCDOEKSTPLQNIIPA----KSLIPAVKITE 293

Qy 301 PIKKEVPAKSDVYCEVCEFLVKEVTKLIDNNKTEKEILDADFCKWCKLPKSLSEBCEQV 360  
 Db 294 ----NPLP--ENNVYCEVCELMISQIEKLDNNTRNINIKOSLEKVLKLPQSVQTKEDI 348

Qy 361 VDTYGSLSILLESVPELVCSMLHLCSGTR--LPALTIVHTVPKDGDFCEVCKLVGYL 419  
 Db 349 IDEYCDPLIELLEQEANPEVICTTLGYCSGRKLNKVKISAEKVAAGDYCPVCRMIRV 408

Qy 420 DRNLEKNSTKQEIILAELEKGSFLPDPYQKQDQFVAEYEPVLIELVEMDPSFVCLIKI 479  
 Db 409 DELLERKATESRIKDFNLICNLFPLDSMQNECSALIKYEYEPFIQLLEBALDPSFCLKL 468

Qy 480 GACPSAKHKLIGTEKCIWGPSYQCONTEAQAACNAVEHCHKRHVN 524  
 Db 469 HLC--QGETVLLGTCKMVGPSYQKVDVETAANCNALEHCHKRHVN 512

RESULT 3  
 Q8UVZ4





```
Qy 123 GEMSRP-GEVCSALNLCESLQKHLAELNHQKLESNKIPELDMTEVVPFMANIPILLYP 181
Db 127 GAQSTDLASVCTALTCEPLQRLHAETTSERPLTQE-----DANEVMAFSLNGALSFHP 181
Qy 182 QDGRSPKQPKDNGVCDQCIQWTDIQTAVRTNSTFVOALVEHVKEECDRLPGMGADIC 241
Db 182 SQMP-----EGAVCHDCVQLISLLQDALESNLTLAEVTQN---QCQSMGFLAALC 230
Qy 242 KNYISQYSEIAIOMMHMPKEICALVGFCDVEKEMPMTQVLPAKVASKNVIIPALELVEP 301
Db 231 ENYTHRFVPAKQTOGLPPEVCRKGGFCERESAHWL-----TRVAADVGVPSLEMP 285
Qy 302 IKHEVPAKSVYCEVCEFLVKEVTKLIDNNKTEKILDAFDMCKSLPKSLSECEQEV 361
Db 286 -RTNELQWLQGTCDVCLNLVQELDKWLVNNSALISHTLERVCTVPEPLVQCCITLV 344
Qy 362 DTYGSSILSLLEVSPELVCSMLHLCGTR-----LPALTVHVTPQKDGFC 410
Db 345 DTSPELVQ-LMSKVTPEKVCETIKLCGSKRRARSISRAVATPSLP--VDEENQGSFCQ 401
Qy 411 VCKLVGLDRNLKNSKQKILAALEKGC--SFLPDPYQKQCD 452
Db 402 GCKELIGH-----VFPESGPQEHQAG-HSECLQRLWLPDVALCD 440
```

## RESULT 7

```
Q8C1N0 PRELIMINARY; PRT; 402 AA.
AC Q8C1N0, 2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Hypothetical saposin A-type domain/saposin type B containing
DE protein.
GN 2310020A21RIK.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Tongue;
RX MEDLINE=22354683; PubMed=12466851;
RA The FANTOM Consortium,
RA the RIKEN Genome Exploration Research Group Phase I & II Team;
RT "Analysis of the mouse transcriptome based on functional annotation of
RL 60,770 full-length cDNAs.";
RL Nature 420:563-573(2002).
DR EMBL; AK009408; BAC25258.1; -
DR MGD; MGI:1924193; 2310020A21RIK.
DR GO; GO:0005764; C:lysosome; IEA.
DR GO; GO:0006665; P:sphingolipid metabolism; IEA.
DR InterPro; IPR003119; Sapa.
DR InterPro; IPR007856; Sapa.1.
DR InterPro; IPR008138; Sapa.2.
DR InterPro; IPR008140; Sapa.sub.
DR InterPro; IPR008373; Saposin.
DR InterPro; IPR008139; SaposinB.
DR Pfam; PF02199; SAPA; 1.
DR Pfam; PF05184; Sapa.1; 3.
DR Pfam; PF03489; Sapa.2; 3.
DR PRINTS; PR01797; Saposin.
DR ProDom; PD001732; Sapa.sub; 2.
DR SMART; SM00162; SAPA; 1.
DR SMART; SM00118; SAPA; 3.
KW Hypothetical protein.
SQ SEQUENCE 402 AA; 44420 MW; E90017CBF4017ED6 CRC64;
```

Query Match 27.7%; Score 772; DB 11; Length 402;  
Best Local Similarity 38.1%; Pred. No. 2.4e-51;  
Matches 154; Conservative 85; Mismatches 129; Indels 36; Gaps 8;

```
Qy 131 VCSALNLCESLQKHLAELNHQKLESNKIPELDMTEVVPFMANIPILLYPQDGRSPKQ 190
Db 13 VCTALTCEPLQRLHAETTSERPLTQE-----DANEVMAFSLNGALSFHPQMP----- 62
Qy 191 PKONGDVCDQCIQWTDIQTAVRTNSTFVOALVEHVKEECDRLPGMGADICNKYISQYSE 250
Db 63 ---EGAVCHDCVQLISLLQDALESNLTLAEVTQN---QCQSMGFLAALCENTYHROFV 116
Qy 251 IAIOMMHMPKEICALVGFCDVEKEMPMTQVLPAKVASKNVIIPALELVEPIKKEHVPAK 310
Db 117 PAKOTLOGLPPEVCRKGGFCERESAHWL-----TRVAADVGVPSLEMP-RTNELQMQ 170
Qy 311 SDVYCEVCEFLVKEVTKLIDNNKTEKILDAFDMCKSLPKSLSECEQEVVDYTGSSILS 370
Db 171 LGLTCDVCLNLVQELDKWLVNNSALISHTLERVCTVPEPLVQCCITLVDPSELVQ 230
Qy 371 ILLEVSPELVCSMLHLCGTR-----LPALTVHVTPQKDGFCVCKKLVGYL 419
Db 231 -LMSKVTPEKVCETIKLCGSKRRARSISRAVATPSLP--VDEENQGSFCQCKRLGMS 287
Qy 420 DRNLKNSKQKILAALEKGCFLPDPYQKQCDQFVAEYEPVLIEILVEMDPSFVCLKI 479
Db 288 SQNLDHKSTKRDILNAPKGGCRILPLPYVMQCNRFVAEYEPVLIESLKFMNPTDLCKKM 347
Qy 480 GAPSAAHKPLLTGTEKICWGSYMCQNTETATAQCNVAVHCKRHVV 523
Db 348 GACHGPKTLLGTQDCVMGSPFWCKSPAAAEMCNALSHCQRLVW 391
```

## RESULT 8

```
Q9Y125 PRELIMINARY; PRT; 953 AA.
AC Q9Y125;
DT 01-NOV-1999 (TrEMBLrel. 12, Created)
DT 01-NOV-1999 (TrEMBLrel. 12, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE SAP-R protein.
GN SAP-R OR BCDA:GH08312 OR CGI2070.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Berkely.
RX MEDLINE=20196006; PubMed=10731132;
RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Ananides P.G., Scher S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Wortman J.R., Vandal M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Niklos G.L.G.,
RA April J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brotter P.,
RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Dou P.L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
RA Foeller C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA Glodok A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
RA Jafari M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Laoko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,
```

```

RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Shue B.C., Siden-Klamas I., Simpson M., Skupeki M.P., Smith T.,
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Wasserman D.A., Weinstein G.M., Weissenbach J.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yeh R.-P., Zaveri J., Zhang S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RA "The genome sequence of Drosophila melanogaster.";
RL Science 287:2185-2195(2000).
RN [2]
RP SEQUENCE FROM N.A.
RA Rubin G.M., Wan K.H., Harvey D., Lewis S.E., Brokstein P., Tsang G.,
RA Agbayani A., Arcania T.T., Baxter E., Blazef R.G., Butenhoff C.,
RA Champe M., Chavez C., Chew M., Doyle C.M., Farfan D.E., Frise E.,
RA Galle R., George R.A., Harris N.L., Hoskins R.A., Evans-Holm M.,
RA Houston K.A., Hummasti S.R., Kim E., Li P., Moehrefi M., Pacleb J.M.,
RA Park S., Sequeira A., Sethi H., Snir E., Svirskas R.R., Weinburg T.,
RA Celnikier S.E.;
RA "Full length Drosophila melanogaster cDNA sequence.";
RL Submitted (APR-1999) to the EMBL/GenBank/DBSJ databases.
DR EMBL; AF003775; AAD57097.1; -.
DR FLYBASE; FBgn000416; Sap-1.
DR GO; GO:0005764; C:lysosome; IEA.
DR GO; GO:0006665; P:sphingolipid metabolism; IEA.
DR InterPro; IPR003119; Sapa.
DR InterPro; IPR007856; SappB_1.
DR InterPro; IPR008138; SappB_2.
DR InterPro; IPR008140; SappB_sub.
DR InterPro; IPR008373; Saposin.
DR InterPro; IPR008139; SaposinB.
DR Pfam; PF02199; SAPA_1.
DR Pfam; PF03489; SappB_1; 6.
DR Pfam; PF03489; SappB_2; 7.
DR PRINTS; PR01797; SAPOSIN.
DR ProDom; PD001732; SappB_sub; 6.
DR SMART; SM00162; SAPA; 1.
DR SMART; SM00118; SAPP; 7.
SQ SEQUENCE 953 AA; 105962 MW; D6CFPD3E9D1502A8 CRC64;

Query Match
Best Local Similarity 19.8%; Score 552; DB 5; Length 953;
Matches 147; Conservative 107; Mismatches 220; Indels 162; Gaps 19;

QY 1 MVALFLASLLGA-ALAGPVLGLKCTRGSAVWQNVKTASDCGAVKCLQTVNWK---P 56
DB 6 LLAVLALCCAGVFAATPLIGSSKCTWGPSYCGFNSKCRATHRCIQTVWETQKVP 65
QY 57 TVKSLPCDICHVVTAAAGMLKONATBEELVYLEKTCDWLPKPNMSASCKEIVDSYLPV 116
DB 66 VDTDSICTCKDMVTQARDQLKSNQTEELKEVFGSGKLIPIKIQKECIKVAADFLEE 125
QY 117 ILDTIKEMSPRGVCSALNICS-----LQKH----- 144
DB 126 LVEALASQNM-PDQVCSAGLGNARSIDELYKNGIQAGLDGTQVQNBDSSEETELAMQPN 184
QY 145 -----LAELNHQKQLESNKIPELDMTEV-----APPMANIPILL----- 179
DB 185 QLSGCGNCLLSRLMHSKFAATDRD---DMVETMLHMGSGLSFSDACANIVITYFNIDYD 241
QY 180 -----YQDGRSPQPK-----DNGD-----VQDCICQMVTDIQ 209
DB 242 HVSKHLTDAVCHVSGVCSARYHQHEEKQPQEQBALVALDAGDDIPCELCEQLVKHLRDVL 301
QY 210 TAVRTNSTFFQALVHVHKECDRLGPGMADICKNYISQYSEAIQ-MMHMQPKICAIY 268
DB 302 VANTTEFEQVMEGCKQ-----SKGFKDECLISVDQYHYHYIETVLSKLDANGACCM 356
QY 269 GFCDE-----VKEMPQTLVPAKASKNVPALDEL---VEPIKKEH----- 306

```

```

Db 357 GICOKNSASSMKDVPIMPILP-----VIEPAQVKITIEKLEKHEKKQIGASEPKFSQ 409
QY 307 -----VP-----AKSDVYCEVCEFLVKEVTKLIDNNKTEKILDAFDKMC 347
Db 410 EILDMLQPLIDHLMGAANPGALVEGGELCTLCYMLHFIQETLATPSTDDEIKHTVENICA 459
QY 348 KLPKSLSEECQEVVDVTGSSILSILEEVSPELVCSMLHLCSTRLPALTVHVTQP----- 403
Db 470 KLPQSVAGQCQNFVEMGVDAVIALVQGLNPRDVCFLMQMCPKNLPKKEDEVENFPQAS 529
QY 404 --XDGFCFEVCKLVGLDNLNLEKNSTKQBIILAALEKGCFLPDYQKQCDQVAVPEPV 461
Db 530 DEQDPPTCPLCLFAVEQAQMKIRDNSKDNIKVNLGLNGLSHLPNEKEECVDFVNTYSNE 599
QY 462 LILILVEWMDPSFVCLKIGACPSAHKPLLGTEKCIW 497
Db 590 LIDMLITDFKPOEICVOLKCPK-----TTVALW 618

RESULT 9
Q8N7T4 PRELIMINARY; PRT; 241 AA.
ID Q8N7T4
AC Q8N7T4
DT 01-OCT-2002 (TReMBLrel. 22, Created)
DT 01-OCT-2002 (TReMBLrel. 22, Last sequence update)
DT 01-OCT-2003 (TReMBLrel. 25, Last annotation update)
DE Hypothetical protein FLJ40379.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Testis;
RA Oshima A., Takahashi-Fujii A., Tanase T., Imose N., Takeuchi K.,
RA Arita M., Musashino K., Yuuki H., Hara H., Sugiyama T., Irie R.,
RA Otsuki T., Sato H., Wakamatsu A., Ishii S., Yamamoto J., Isono Y.,
RA Kawai-Hio Y., Saito K., Nishikawa T., Kimura K., Yamashita H.,
RA Matsuo K., Nakamura Y., Sekine M., Kikuchi H., Kanda K., Wagatsuma M.,
RA Murakawa K., Kanehori K., Sugiyama A., Kawakami B., Suzuki Y.,
RA Sugano S., Nagahari K., Masuho Y., Nagai K., Isogai T.;
RT "NEO human cDNA sequencing project.";
RL Submitted (JUL-2002) to the EMBL/GenBank/DBSJ databases.
DR EMBL; AK097698; BAC05143.1; -.
DR GO; GO:0005764; C:lysosome; IEA.
DR GO; GO:0006665; P:sphingolipid metabolism; IEA.
DR InterPro; IPR003119; Sapa.
DR InterPro; IPR007856; SappB_1.
DR InterPro; IPR008138; SappB_2.
DR InterPro; IPR008140; SappB_sub.
DR InterPro; IPR008373; Saposin.
DR InterPro; IPR008139; SaposinB.
DR Pfam; PF02199; SAPA_1.
DR Pfam; PF03489; SappB_1; 2.
DR Pfam; PF03489; SappB_2; 2.
DR ProDom; PD001732; SappB_sub; 2.
DR SMART; SM00162; SAPP; 1.
DR SMART; SM00118; SAPP; 2.
DR Hypothetical protein.
SQ SEQUENCE 241 AA; 26719 MW; BE29EE4F6FEFE6AB CRC64;

Query Match
Best Local Similarity 19.5%; Score 545; DB 4; Length 241;
Matches 107; Conservative 45; Mismatches 76; Indels 12; Gaps 4;

QY 293 IPALVEPIKKEHVEPAKSDVYCEVCEFLVKEVTKLIDNNKTEKILDAFDKMCXLPKS 352
DB 4 VPSLELGLPKQSEMOMKAGVTCEVCMVQKLDHLMNSSELMTHALERVCVSPAS 63
QY 353 LSEECQEVVDVTGSSILSILEEVSPELVCSMLHLCSTRLPALTVH-----VTQPK--- 404
DB 64 ITKEIILVDVTYSPLVQ-LVAKITPEKVCFKIRLC-GNRRRARAVHDAIYVSPWDMA 121

```



```
QY 405 --DGGFCVCKLVGYLDRLNLEKNTKQETLAALAEKGCFLPDPYQKQCQDFVAEYEPVL 462
DB 122 ENQGSFNGCKRLLTSSHNLESKSTKRDILVAFKGCSTLPLPYMIQCKHFVQYEPVL 181
QY 463 IEILVEWDSFVCLKIGACSAHKPLIGTEKCTWGSYWCNVTETAAQCNAVCHCKRHV 522
DB 182 IESLKMDMPVAVCKVYGGACHGPRPTLLGTDTQCALGSPFWCRSQEAARKLCNAVQCHQHV 241

RESULT 10
P79254
ID P79254 PRELIMINARY; PRT; 121 AA.
AC P79254;
DT 01-MAY-1997 (TrEMBLrel. 03, Created)
DT 01-MAY-1997 (TrEMBLrel. 03, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Sulfated glycoprotein-1/SGP-1 (Fragment).
OS Ovis aries (Sheep).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovidae; Caprinae; Ovis.
OX NCBI_TaxID=9940;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=96249303; PubMed=8848570;
RA Spencer T.E., Graf G.H., Bazer F.W.;
RT "Sulfated glycoprotein-1 (SGP-1) expression in ovine endometrium
during the oestrous cycle and early pregnancy.";
RL Reprod. Fertil. Dev. 7:1053-1060(1995).
DR EMBL; S82555; AAD14405.1; -.
DR GO; GO:0005764; C:lysosome; IEA.
DR InterPro; IPR008138; SGP_2.
DR InterPro; IPR008373; Saposin.
DR InterPro; IPR008139; SaposinB.
DR Pfam; PF03489; SGP_2; 1.
DR PRINTS; PR01797; SAPOSIN.
DR SMART; SM00118; SAPB; 1.
FT NON TER 1
SQ SEQUENCE 121 AA; 13604 MW; 4F0F5A6EB83D0C9A CRC64;

Query Match 18.5%; Score 517; DB 6; Length 121;
Best Local Similarity 77.5%; Pred. No. 2.5e-32;
Matches 94; Conservative 18; Mismatches 9; Indels 0; Gaps 0;

QY 205 VTDIQTAVRTNSTFVQALVEHVKECDRLGPGMADICKNYISOYSEIAIQMMHMQPKEI 264
DB 1 VTDIQTAVRTNSTFVQGLVDHVKECDRLGPGMADICKNYLAQYSEIAVQMMHMQPKEI 60

QY 265 CALVGFCDCEKMPQTLVPAKVASKNVI PALELVEPIKKEHVPKSDVYCEVCEFLVKE 324
DB 61 SALVGFCDCEKESPARTLDPAKEASKNVLPALELTEPNKQEI PAQTLIFCQVCQFVRE 120

QY 325 V 325
DB 121 V 121

RESULT 11
O15997
ID O15997 PRELIMINARY; PRT; 965 AA.
AC O15997;
DT 01-JAN-1998 (TrEMBLrel. 05, Created)
DT 01-JAN-1998 (TrEMBLrel. 05, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE BmPi09.
OS Bombyx mori (Silk moth).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Lepidoptera; Glossata; Ditrysia; Bombycoidea;
OC Bombycidae; Bombyx.
OX NCBI_TaxID=7051;
RN [1]
```

```
RP SEQUENCE FROM N.A.
RX MEDLINE=98278844; PubMed=9611271;
RA Tambunan J., Chang P.-K., Li H., Natori M.;
RT "Molecular cloning of a cDNA encoding a silkworm protein which
contains the conserved BH regions of Bcl-2 family proteins.";
RL Gene 212:287-293(1998).
DR EMBL; AB008449; BAA23126.1; -.
DR PIR; T00207; T00207.
DR GO; GO:0005764; C:lysosome; IEA.
DR InterPro; IPR003119; Sapa.
DR InterPro; IPR007856; SGP_1.
DR InterPro; IPR008138; SGP_2.
DR InterPro; IPR008140; SGP_sub.
DR InterPro; IPR008373; Saposin.
DR InterPro; IPR008139; SaposinB.
DR Pfam; PF02139; SAPA; 2.
DR Pfam; PF05184; SGP_1; 6.
DR Pfam; PF03489; SGP_2; 7.
DR PRINTS; PR01797; SAPOSIN.
DR ProDom; PD001732; SGP_sub; 6.
DR SMART; SM00162; SAPA; 2.
DR SMART; SM00118; SAPB; 7.
SQ SEQUENCE 965 AA; 108825 MW; F41A7BEE7F626078 CRC64;

Query Match 18.3%; Score 511; DB 5; Length 965;
Best Local Similarity 24.6%; Pred. No. 1e-30;
Matches 139; Conservative 104; Mismatches 229; Indels 94; Gaps 18;

QY 19 VLGLKECTRGSAVWCVNTASDCGAVKHCLOTVWNK---PTVKSLPCDICKVWVTAAGD 75
DB 170 LLGSRCTGWSYWCNFSFGRECNPATPHCINRWSKMTFPEDNDNICQICLDMVKOARD 229

QY 76 MLKDNATEREILVLEKTDMLPKPMNSASCKEIVDSYLPVILDIKGENSRPCEVCAL 135
DB 230 QLOSNETQDEIKVEFEGSCKLIPIKFAECMKLADEFVVELIETLASEMN-POAVCSVA 288

QY 136 NLCS--LQKHLAEHLNQKLES-----NKPELDMTEVYVAPFMA----- 173
DB 289 GLCNNAKIDRLLDVYNAQRELGRAGCYNQCTGVWRKKFDTEKYEDFLVGLQVCRNMS 348

QY 174 ---NIPLLLYP-----QDG-----PRSKPQPKDNGD 196
DB 349 LSDSCSNLIPKYYENILEAVKDLNPEGICHVSGQCSYKFNHDEFTFPQMVQYSATDD 408

QY 197 V-QDCICIQMTDITQTAVRTNSTFVQALVEHVKECDRLGPGMADICKNYISOYSEIAIQM 255
DB 409 VPCEFCQLVKHRLDVLVANTTELE-FYKVLQGLCKQTGK-FKDECIHLAEQYYPVLYNF 466

QY 256 MM-HMQPKETCALVGFCDCEKMPQTLV-----PAKVA-----SKNVIP 294
DB 467 LVSDLPKPAETCKMIGICGNLTSAPIPLVARELVVQPKLIGAEESKIARVPLAKQMEP 526

QY 295 ALELVE--PTKKEHVPK-KSDVYCEVCEFLVKEVTKLIDNNKTEKEILDADKMKCKLPK 351
DB 527 ASAAVSPLPLERMFVAAPQSKAACAFQCYFLHYLVQLSDTRTEDKVKAAVQEAACDALPD 586

QY 352 SLSEECQEVVDYTGSSITSLTLLLEVPSELVCSMLHLCSTGLPALTVHVTPQKDGGECEV 411
DB 587 ALNGECKBFVTQYGSVAIVALLVQEI DPASVCPALQICPQTE-EIRRVDMNSEKN--CPL 643

QY 412 CKKLGVYLDRLNLEKNTKQETLAALAEKGCFLPDPYQKQCQDFVAEYEPVLIELVWMD 471
DB 644 CLFAVEQLSVLKNRSEENIRKALDGLCTRLSKQLQSECIQDFVDTYSSQLVEMLVADMN 703

QY 472 PSFVCLKIGAC-PSAHKPLIGTEKCI 496
DB 704 AKEICVFLKLCRDQLDPLKLTSSSI 729

RESULT 12
Q8IMH4 PRELIMINARY; PRT; 876 AA.
ID Q8IMH4
```

AC Q8IMH4;  
 DT 01-MAR-2003 (TrEMBLrel. 23, Created)  
 DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)  
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
 DE CG12070-PB.  
 DE SAP-R.  
 GN Drosophila melanogaster (Fruit fly).  
 OS Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;  
 OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;  
 OC Ephydroidea; Drosophilidae; Drosophila.  
 OX NCBI\_TaxID=7227;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=20196006; PubMed=10731132;  
 RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,  
 RA Amanatides P.G., Scher S.E., Li P.W., Hoskins R.A., Gallie R.F.,  
 RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,  
 RA Sutton G.G., Mortman J.R., Yandell M.D., Zhang Q., Chen L.X.,  
 RA Brandon R.C., Rogers Y.H., Blazey R.G., Champagne M., Pfeiffer B.D.,  
 RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Gabor G.L.,  
 RA Abril J.F., Agbayani A., An H.J., Andrews-Pfannkuch C., Baldwin D.,  
 RA Balles R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,  
 RA Beeson K.Y., Boman P.V., Berman B.P., Bhandari D., Bolehakov S.,  
 RA Borkova D., Borchan M.R., Bouck J., Brokstein P., Brothier P.,  
 RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,  
 RA Cherry J.M., Cawley S., Dahke C., Davenport L.B., Davies P.,  
 RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,  
 RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,  
 RA Durbin K.J., Evangelista C.C., Ferraz C., Ferrieria S., Fleischmann W.,  
 RA Foster C., Gabriellian A.E., Garg N.S., Galbart W.M., Glaesner K.,  
 RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,  
 RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,  
 RA Hostin D., Houston K.A., Howland T.J., Wei M.H., Ibegwam C.,  
 RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,  
 RA Kimble B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,  
 RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,  
 RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,  
 RA Merkulov G., Milehina N.V., Mobarry C., Morris J., Moshrefi A.,  
 RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,  
 RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacled J.M.,  
 RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,  
 RA Reinert K., Remington K., Saunders R.D., Scheeler F., Shen H.,  
 RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,  
 RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,  
 RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,  
 RA Wang Z.Y., Wasserman D.A., Weinstock G.M., Weissbach J.,  
 RA Williams S.M., Woodger, Worley K.C., Wu D., Yang S., Yao Q.A., Ye J.,  
 RA Yeh R.F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,  
 RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,  
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.,  
 RT "The genome sequence of Drosophila melanogaster."  
 RL Science 287:2185-2195(2000).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RA Celniker S.E., Adams M.D., Kronmiller B., Wan K.H., Holt R.A.,  
 RA Evans C.A., Gocayne J.D., Amanatides P.G., Brandon R.C., Rogers Y.,  
 RA Banzon J., An H., Baldwin D., Banzon J., Beeson K.Y., Busam D.A.,  
 RA Carlson J.W., Center A., Champagne M., Davenport L.B., Dietz S.M.,  
 RA Dodson K., Dorsett V., Doup L.E., Doyle C., Dresnek D., Farfan D.,  
 RA Ferreira S., Frise E., Galle R.F., Garg N.S., George R.A.,  
 RA Gonzalez M., Houck J., Hoskins R.A., Hostin D., Howland T.J.,  
 RA Ibegwam C., Jalali M., Kruse D., Li P., Mattei B., Moshrefi A.,  
 RA McIntosh T.C., Moy M., Murphy B., Nelson C., Nelson K.A., Nunoo J.,  
 RA Pacled J.C., Paragov S., Park S., Patel S., Pfeiffer B.,  
 RA Phouanavong S., Pittman G.S., Puri V., Richards B., Scheeler F.,  
 RA Stapleton M., Strong R., Svirskas R., Tector C., Tyler D.,  
 RA Williams S.M., Zaveri J.S., Smith H.O., Venter J.C., Rubin G.M.;  
 RT "Sequencing of Drosophila melanogaster genome."  
 RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RA Miera S., Crosby M.A., Matthews B.B., Bayraktaroglu L., Campbell K.,  
 RA Hradecky P., Huang Y., Kaminker J.S., Prochnik S.E., Smith C.D.,

RA Tupy J.L., Bergman C.M., Berman B.P., Carlson J.W., Celniker S.E.,  
 RA Clamp M.E., Drysdale R.A., Emmert D., Frise E., de Grey A.D.N.J.,  
 RA Harris N.L., Kronmiller B., Marshall B., Millburn G.H., Richter J.,  
 RA Russo S., Searle S.M.J., Smith E., Shu S., Smutniak F.,  
 RA Whitfield E.J., Ashburner M., Gelbart W.M., Rubin G.M., Mungall C.J.,  
 RA Lewis S.E.;  
 RT "Annotation of Drosophila melanogaster genome."  
 RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.  
 RN [4]  
 RP SEQUENCE FROM N.A.  
 RA FlyBase;  
 RL Submitted (SEP-2002) to the EMBL/GenBank/DBJ databases.  
 RN [5]  
 RP SEQUENCE FROM N.A.  
 RA FlyBase;  
 RL Submitted (FEB-2003) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AE003775; AN14261.2; -  
 DR GO; GO:0005764; C:lysosome; IEA.  
 DR GO; GO:0005489; F:electron transporter activity; IEA.  
 DR GO; GO:0006118; P:electron transport; IEA.  
 DR GO; GO:0006665; P:phospholipid metabolism; IEA.  
 DR InterPro; IPR000345; CytC heme\_BS.  
 DR InterPro; IPR007856; SapB\_1.  
 DR InterPro; IPR008138; SapB\_2.  
 DR InterPro; IPR008140; SapB\_sub.  
 DR InterPro; IPR008373; Saposin.  
 DR Pfam; PF05184; SapB\_1; 3.  
 DR Pfam; PF03489; SapB\_2; 4.  
 DR PRINTS; PR01797; SAPOSIN.  
 DR ProDom; PD001732; SapB\_sub; 2.  
 DR SMART; SM00118; SAPB; 7.  
 DR PROSITE; PS00190; CYTOCHROME\_C; 1.  
 DR SEQUENCE 876 AA; 97617 MW; 767E16D35ACF52DB CRC64;  
 Query Match 15.5%; Score 432.5; DB 5; Length 876;  
 Best Local Similarity 23.8%; Pred No. le-24;  
 Matches 117; Conservative 91; Mismatches 178; Indels 105; Gaps 17;  
 QY 61 LPDCKDVVTAAGDMLKDNATEEILVLEKTDMLPKNNASCKEIVDSYLPVLDI 120  
 DB 208 IPCECEQLVRLDLVANTTETETEFKQVMEGCK--QSKGFKDECLISVDQYHYIYET 265  
 QY 121 IKGESRRPQVCVCSALNLC-----ESLQKHAEINLHQ 151  
 DB 266 LVSKLDANG-ACCMIGICQKNASMKDVPIMPLLPVIEPAQVKITTEKLEK 319  
 QY 152 KQLESNKIPELDMTVEVWAFPMANIPLLLYPODGPRSKPQP---KNGDGVQDCIQWVTDI 208  
 DB 320 KQLGASE-PKFSQOEILD-----MQLPIDLHMGANFGALVEGELCTLCYMLHFI 370  
 QY 209 QTAVRTNSTFYQALVEHVKEE-CDRLGPGMADICKNVIQYSEIATQMMH-MQPKICA 266  
 DB 371 QETLATPSTDD--IKHTVENICAKLPESGVAGQCRNEVEMYGDAVIALLVQGLNPRDVP 428  
 QY 267 LVGFCDEKEMPOTLVPAKVASKNVIPALELVEPIKKEHVPAKSDVYCEVCFYKEVT 326  
 DB 429 LMQCP--KNLPK-----EDVEVFNPPQASDEQDP-----PT-----CPCLFLAVEQAQ 471  
 QY 327 KLINNKTEKILDAFKMCKPKLSSECOBWDVTGSSILSILEEVSPELVCSMLH 386  
 DB 472 MKIRDNKSCKNIKKVNLGLCSLHPNEIKBCVDFVNTYSNELIDMLITDPKPOEICVOLK 531  
 QY 387 LCSGT--RLPALTVHVTQPKDG-----GF-----CEV 411  
 DB 532 LCPKTTVALWDLRLISLEDDVDGDKSSSEISFNDESLEELPPQLAFDPGFTAAPNCLI 591  
 QY 412 CKKLGVGLDNLNLEKNSTKQBIKLAALAEKGCSEFLPDYKQCDQFVAEYEPVLIEILVEVMD 471  
 DB 592 CEELVKTLEKMGKHPTRDSIKHILESODRMKPNWTKCHKVIDKYGDKIADLLKEND 651  
 QY 472 PSFVCLKIGAC 482

```
Db 652 PKLICTELGMC 662
RESULT 13
O35489
ID O35489 PRELIMINARY; PRT; 378 AA.
AC O35489;
DT 01-JAN-1998 (TrEMBLrel. 05, Created)
DT 01-JAN-1998 (TrEMBLrel. 05, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Surfactant protein-B.
OS Cavia porcellus (Guinea pig).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Hystricognathi; Caviidae; Cavia.
OX NCBI_TaxID=10141;
RN [1]
RP SEQUENCE FROM N.A.
RA Bingle C.D., Yuan H.T., Gowan S.;
RT "Guinea pig surfactant protein-B is differentially polyadenylated.";
RL Submitted (NOV-1997) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF033190; AAB87086.1; -.
DR HSSP; P07988; 1DFW.
DR GO; GO:0005764; C:lysosome; IEA.
DR GO; GO:0007585; P:respiratory gaseous exchange; IEA.
DR GO; GO:0006665; P:sphingolipid metabolism; IEA.
DR InterPro; IPR003119; SApA.
DR InterPro; IPR007856; SApB_1.
DR InterPro; IPR008138; SApB_2.
DR InterPro; IPR008140; SApB_sub.
DR InterPro; IPR008373; Saposin.
DR InterPro; IPR008139; SaposinB.
DR InterPro; IPR008137; Surfactant_B.
DR Pfam; PF02199; SApA; 1.
DR Pfam; PF05184; SApB_1; 2.
DR Pfam; PF03489; SApB_2; 3.
DR PRINTS; PR01797; SAPOSIN.
DR PRODOM; PD001732; SApB_sub; 1.
DR PRODOM; PD008002; Surfactant_B; 1.
DR SMART; SM00162; SApA; 1.
DR SMART; SM00118; SApB; 3.
SQ SEQUENCE 378 AA; 41677 MW; E3DA2E237ED401DA CRC64;

Query Match 12.5%; Score 349.5; DB 11; Length 378;
Best Local Similarity 22.9%; Pred. No. 9.2e-19;
Matches 113; Conservative 67; Mismatches 175; Indels 139; Gaps 15;

Qy 4 LFLASLIGALAGVLGKECTRGSAVNCQVKTASDCGAVKICLOTVWVKPTVKSIPC 63
Db LVLLPMLCGPAAQAQTSSSLACSQGPKFCWCSLEQALQCRAIGHCLQEVGHVGADDL-C 68

Qy 64 DICKDVVTAAGDMLKDNATEEEILVLEKTCDWLPKPNMSCKEIVDSYLPVLIDTIKG 123
Db QECEDIVRLTYMAKEVIFQTKIRFLERECADVLPKLLVPRCHSVLETYPFLVIDYFQS 128

Qy 124 EMSRPGVCSALNTCESLQKHLAEINHQLESNKIPELDMTEVVAPPMANIPLLLYPD 183
Db 129 HIT-PKTCWNLGLCQPRQ-----PD----- 148

Qy 184 GPRSKPQKNGDVQCDCIQMVTDIQTAVRTNSTFVQALVEHVEECDRLPGHADICKN 243
Db 149 -----POPE----- 161

Qy 244 YISQVSEITAIQMMHMQPEICALGVCFDEVKEMPQTLVPKAV-ASKNVIIPALELVEPI 302
Db 162 SLSETSP-----DLLVPPRLPRLAUSVMPGPH-TQDL 191

Qy 303 KKEHVPKSDVYCEVCEFLVKEVTKLIDNNKTEKILD-AFDKMSCKLPKSLSECEQEV 361
Db 192 SEQQFPPIPLP-YCRCLKTLKRVQAMI-----PKGVLAMAVAQCHVVPVAVGICQCLA 245

Qy 362 DTYGSSILSILEEVSPELVCSMLHLCS-----GTRLPALTVHTVTPQKDGFCFCEVKLV 416
Db 246 ERYTVLLDALLSHLLPOLVGLVLRCSMDNSAGILVLPALSLSEGLPQDSECDLCISVT 305

Qy 417 GYLDRNLERNKSTKOEILAALEKGCFLPDPYQKQ-CDQFVAEYEPVLIEILVEMDSFV 475
Db 306 -----TRAWNSSELARPOAMHQAC--LSSPLDRQCKQFVEQYTFQLLALPRGGDPRTT 358

Qy 476 CLKIGACPSAHKPL 489
Db 359 QOALGVCRGAKNPL 372

RESULT 14
Q95X02 PRELIMINARY; PRT; 458 AA.
ID Q95X02
AC Q95X02;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Naegleria fowleri.
GN NP-B.
OS Naegleria fowleri.
OC Eukaryota; Heterolobosea; Schizopyrenida; Vahlkampfiidae; Naegleria.
OX NCBI_TaxID=5763;
RN [1]
RP SEQUENCE FROM N.A.
RA Nickel R., Leippe M.;
RT "Pore-forming peptides of Naegleria fowleri.";
RL Submitted (OCT-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF196309; AAL01158.1; -.
DR GO; GO:0005764; C:lysosome; IEA.
DR GO; GO:0006665; P:sphingolipid metabolism; IEA.
DR InterPro; IPR007856; SApB_1.
DR InterPro; IPR008138; SApB_2.
DR InterPro; IPR008140; SApB_sub.
DR InterPro; IPR008173; Saposin.
DR InterPro; IPR008139; SaposinB.
DR Pfam; PF05184; SApB_1; 5.
DR Pfam; PF03489; SApB_2; 5.
DR PRINTS; PR01797; SAPOSIN.
DR PRODOM; PD001732; SApB_sub; 4.
DR SMART; SM00118; SApB; 5.
FT NON TER 1
SQ SEQUENCE 458 AA; 50452 MW; 433F60CCD6E4C904 CRC64;

Query Match 12.3%; Score 343.5; DB 5; Length 458;
Best Local Similarity 25.6%; Pred. No. 3.4e-18;
Matches 115; Conservative 69; Mismatches 186; Indels 79; Gaps 16;

Qy 42 CGAVKHC---LQTVWVKPTVKSIPCDICDQVVTAAAGDMLKDNATEEEILVLEKTCDWLP 98
Db 69 CSQVKICNGSAMAAVAPKAENSGICNNMCQLLVTOVENWVESNDTMTLEKLEQVCSVIP 128

Qy 99 KPNMSCKEIVDSYLPVLIDTIKEMSRPG-EVCSALNLCESLQKHLAEINHQLESN 157
Db 129 -GOVSALCTVAEYOYLFIFHQV--EKQFPALTICQDVHLCSAQ----- 170

Qy 158 KIPELDMTEVVAPPMANIPLLLYPDGPRSKPQKNGDVQCDCIQMVTDIQTAVRTNST 217
Db 171 -----AAPV-----CQQAELCPICKAAVGLTKI--NNV 201

Qy 218 FVQALVHVEECDRLPGHADICKNVIQSYSEITAIQMMHMQPEICALGVCFDEVKEM 277
Db 202 DVNAVKQLEFACSF--QVPDCCQIVDRAAQAIA-QDLQTEDAQTCSTV-----VDVC 252

Qy 278 PMQTLVPAKAVSKNVIIPALELVEPIKKEHVPKSDVYCEVCEFLVKEVTKLIDNNKTEKE 337
Db 253 PKQQVW-----TFNPFKKF-LEAKDSKYCPTCLQITKYLEDLIVSDITVNE 297

Qy 338 ILDAFDKMSCKLPKSLSECEQEVVDYTGSSITLSILLEVSPELVCSMLHLCSGTRLPALT 397
Db 298 IIKLADAGCARL-GALESCLKKFFVLAVDELKLLLEKLPQKVCSTLKMCDAAELKLA 356

Qy 398 VHTVTPQKDGFCFCEVKLVGLVLRNLEKNSTKQSBILAALKEGC-SFLPDPYQKQCDQFVA 456
Db 398 VHTVTPQKDGFCFCEVKLVGLVLRNLEKNSTKQSBILAALKEGC-SFLPDPYQKQCDQFVA 456
```

357 L-APQAADGTMCLACEYVISVADNWLNIANNTOOSVKNTLTKVQCFEVPFSIYQSQCIALVN 415

457 EYEPVLIIEL-VEVMDPSPFVCLKIGACPS 484

416 QYEAQLIQLFESKVFNPQTVCRAIGVCSS 444

RESULT 15

Q9BKM1 PRELIMINARY; PRT; 484 AA.

AC Q9BKM1;

DT 01-JUN-2001 (TrEMBLrel. 17, Created)

DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)

DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)

DE Naegleria fowleri. 25, Last annotation update)

GN PRONP-B.

OS Naegleria fowleri.

OC Eukaryota; Heterolobosea; Schizopyrenida; Vahlkampfiidae; Naegleria.

OX NCBI\_TaxID=5763;

RN [1]\_TaxID=5763;

RP SEQUENCE FROM N.A.

RA Nickel R., Benkert C., Jacobs T., Marti T., Marciano-Cabral F.,

RA Lelpe M.;

RT "pore-forming peptides of Naegleria fowleri. ";

RL Submitted (MAY-1399) to the EMBL/GenBank/DDAJ databases.

DR EMBL; AF154047; AAK21659.1; -

DR GO; GO:0005764; C:lysosome; IEA.

DR GO; GO:0006665; P:sphingolipid metabolism; IEA.

DR InterPro; IPR007856; StpB.1.

DR InterPro; IPR008138; StpB.2.

DR InterPro; IPR008140; StpB sub.

DR InterPro; IPR008373; Saposin.

DR InterPro; IPR008139; SaposinB.

DR Pfam; PF05184; StpB.1; 5.

DR Pfam; PF03489; StpB.2; 5.

DR PRINTS; PR01797; SAPOSIN.

DR Prodom; PD001732; StpB sub; 4.

DR SMART; SM00118; StpB; 5.

SQ SEQUENCE 484 AA; 53237 MW; 63DA3AAE7CA578B0 CRC64;

Query Match 12.3%; Score 343.5; DB 5; Length 484;

Best Local Similarity 25.6%; Pred. No. 3.6e-18;

Matches 115; Conservative 69; Mismatches 186; Indels 79; Gaps 16;

QY 42 CGAVKHC---LQTVNNKPTVKSLPCDICKDQVVTAAAGDMLKONATEEELVYLKTCDWLP 98

DB 95 CSQVKICNGSAMA VAAPKAENGSGICNNCQLLVTVQENWVESNTITMLEKLEQVCSVIP 154

QY 99 KPNMSASKEIVDSYLPVILDIKGENSRPG-EVCSALNLCESLQKHLAEINHQKLESN 157

DB 155 -GQVSALCTVAVQYLPFIHQV--EKQFPALATICQDVHLCSAQ-----196

QY 158 KIPELQDTEYVAVFMANIPILLPLYPQDPRSKPQKNGVDVQCDCIQMVTDTQTAVTNST 217

DB 197 -----AAPVW-----QQQAAELCPICKAAGVGLTKI--NNV 227

QY 218 FVQALVHVKEBCDRIGPGMADICKNYISQYSEIAIQMMHMQPEIKALVGFCDVEKEM 277

DB 228 DVNAVXQQLFEAFCSFF---QVPDCQIQVDKAAQIA-QDLQTEDAQITCSTV-----VDVC 278

QY 278 PMQTLVPAKVAKNVIPALELVEPIKHEVPAKSDVYCEVCFVLKEVTYKLIDNNKTEKE 337

DB 279 PKQOVV-----TFNPFKKF-LEAKDSKYCPTCIQITKYLEDLIVSDITVNE 323

QY 338 ILDAFIMCKSKLPKSLSEECQEVVDTYGSSILSLLEEVSPBELVCSMLHLCSGTRLPALT 397

DB 324 ITKLADAGCARL-GALESICKKFPVLADELKLLLEKLPQKVCSTLKWCDAAELLKLA 382

QY 398 VHTVQPKDGFCEVCKKLVGYLDRLNLEKNSTQKELIAALEKCG-SFLPDPYQKCDQFVA 456

DB 383 L-APQAADGTMCLACEYVISVADNWLNIANNTOOSVKNTLTKVQCFEVPFSIYQSQCIALVN 441

QY 457 EYEPVLIIEL-VEVMDPSPFVCLKIGACPS 484

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 5, 2004, 13:16:48 ; Search time 12.9383 Seconds  
(without alignments)  
2108.841 Million cell updates/sec

Title: US-09-743-684A-1  
Perfect score: 2789  
Sequence: 1 MYALFLILASLLGALAGPVL.....NTETAAQCNAVECHRRHVN 524

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 141681 seqs, 52070155 residues

Total number of hits satisfying chosen parameters: 141681

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : SwissProt\_42.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query Match	Score	Length	ID	Description
1	2789	100.0	524	1	SAP_HUMAN
2	2425.5	87.0	525	1	SAP_BOVIN
3	1996	71.6	554	1	SAP_RAT
4	1938.5	69.5	557	1	SAP_MOUSE
5	1679.5	60.2	518	1	SAP_CHICK
6	393	14.1	80	1	SAP_PIG
7	382	13.7	377	1	PSPB_MOUSE
8	379	13.6	376	1	PSPB_RAT
9	349.5	12.5	381	1	PSPB_HUMAN
10	310	11.1	363	1	PSPB_CANFA
11	296.5	10.6	370	1	PSPB_RABIT
12	261	9.4	81	1	SAP_CAVPO
13	173.5	6.2	513	1	ASPR_CUCPE
14	158.5	5.7	508	1	ASPR_HORVU
15	154.5	5.5	473	1	CYPI_CYNCA
16	154	5.5	496	1	ASPR_ORYSA
17	143.5	5.1	1175	1	GLG1_MOUSE
18	142.5	5.1	509	1	APR1_ORYSA
19	138.5	5.0	1171	1	GLG1_RAT
20	137.5	5.0	1179	1	GLG1_HUMAN
21	137.5	4.9	1160	1	GLG1_CRIGR
22	137.5	4.6	1142	1	GLG1_CHICK
23	120	4.3	578	1	EZRA_ENTFA
24	119	4.3	975	1	SEC8_MOUSE
25	118	4.2	857	1	I2C1_HUMAN
26	117	4.2	857	1	I2C1_MOUSE
27	116.5	4.2	1216	1	P1B1_MOUSE
28	116	4.2	975	1	SEC8_RAT
29	114.5	4.1	971	1	RECK_HUMAN
30	114.5	4.1	1216	1	P1B1_RAT
31	114	4.1	984	1	NOF_DROME
32	113.5	4.1	974	1	SEC8_HUMAN
33	113	4.1	418	1	TEK1_HUMAN

RESULT 1

ID	SAP_HUMAN	STANDARD;	PRT;	524 AA.
AC	P07602; P07292; P15793; P78538; P78541; P78546; P78547; P78558;			
AC	Q92739; Q92740; Q92741; Q92742;			
DT	01-APR-1988 (Rel. 07, Created)			
DT	01-APR-1990 (Rel. 14, Last annotation update)			
DE	Proactivator polypeptide precursor (Contains: Saposin A (Protein A); Saposin B (Sphingolipid activator protein 1) (SAP-1) (Cerebroside sulfate activator) (CSAct) (Dispersin) (Sulfatide/GMI activator); Saposin C (Co-beta-glucosidase) (Al activator) (Glucosylceramidase activator) (Sphingolipid activator protein 2) (SAP-2); Saposin D (Protein C) (Component C)).			
GN	PSAP.			
OS	Homo sapiens (Human).			
OC	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.			
OX	NCBI_TaxID=9606;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	TISSUE=Liver;			
RX	MEDLINE=90129043; PubMed=25151510;			
RA	Rorman E.G., Grabowski G.A.;			
RT	"Molecular cloning of a human co-beta-glucosidase cDNA: evidence that four sphingolipid hydrolase activator proteins are encoded by single genes in humans and rats.";			
RL	Genomics 5:486-492(1989).			
RN	[2]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=89255151; PubMed=2498298;			
RA	Nakano T., Sandhoff K., Stuenkel J., Christomanou H., Suzuki K.;			
RT	"Structure of full-length cDNA coding for sulfatide activator, a Co-beta-glucosidase and two other homologous proteins: two alternate forms of the sulfatide activator.";			
RL	J. Biochem. 105:152-154(1989).			
RN	[3]			
RP	SEQUENCE FROM N.A. (ISOFORM SAP-MU-0).			
RC	TISSUE=Brain, Eye, and Skin;			
RX	MEDLINE=223388257; PubMed=12477932;			
RA	Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G., Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D., Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K., Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Heide F., Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L., Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E., Brownstein M.J., Udutin T.B., Toshiyuki S., Carninci P., Prange C., Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J., Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H., Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W., Villalón D.K., Muzny K.C., Sodergren E.J., Lu X., Gibbs R.A., Fahy J., Helton E., Kettelman M., Madan A., Rodrigues S., Sanchez A., Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G., Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C., Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,			

ALIGNMENTS

34	113	4.1	1102	1	MYSC_CHICK
35	113	4.1	2869	1	RBPI_PLAVB
36	112	4.0	1184	1	PBL2_HUMAN
37	112	4.0	1216	1	P1B1_BOVIN
38	111.5	4.0	8545	1	Q9N4M4 CAENORHABDI
39	110.5	4.0	1971	1	Q9WUJ9 MUS MUSCULU
40	110.5	4.0	3911	1	Q99996 H A-KINASE
41	110	3.9	476	1	P53789 ORYCTOLAGUS
42	110	3.9	1216	1	P1B1_HUMAN
43	109	3.9	861	1	I2C4_HUMAN
44	109	3.9	1557	1	DVAL_DICVI
45	108.5	3.9	3674	1	SPCR_HUMAN

P29616	gallus gall
Q00798	plasmodium
P98095	homo sapien
P10894	bos taurus
Q9N4M4	caenorhabdi
Q9WUJ9	mus musculu
Q99996	h a-kinase
P53789	oryctolagus
Q9Nq66	homo sapien
Q9hck5	homo sapien
Q24702	dictyocaulu
Q9nrc6	homo sapien

RA Schnerch A., Schein J.E., Jones S.J.M., Merri M.A.;  
RT "Generation and initial analysis of more than 15,000 full-length  
RL human and mouse cDNA sequences."  
RP Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).  
[4]  
RA SEQUENCE OF 59-125 AND 304-513 FROM N.A.  
RP TISSUE=Brain;  
RX MEDLINE=91192146; PubMed=2013321;  
RA Holtschmidt H., Sandhoff K., Fuerst W., Kwon H.Y., Schnabel D.,  
SU Suzuki K.;  
RT "The organization of the gene for the human cerebroside sulfate  
RT activator protein."  
RL FEBS Lett. 280:267-270(1991).  
[5]  
RA SEQUENCE OF 164-524 FROM N.A.  
RP MEDLINE=88068647; PubMed=2825202;  
RA Dewji N.N., Wenger D.A., O'Brien J.S.;  
RT "Nucleotide sequence of cloned cDNA for human sphingolipid activator  
RT protein 1 precursor."  
RL Proc. Natl. Acad. Sci. U.S.A. 84:8652-8656(1987).  
[6]  
RA PARTIAL SEQUENCE OF 60-142.  
RP MEDLINE=89240739; PubMed=2717620;  
RX Morimoto S., Martin B.M., Yamamoto Y., Kretz K.A., O'Brien J.S.,  
RA Kishimoto Y.;  
RT "Saposin A: second cerebroside activator protein."  
RL Proc. Natl. Acad. Sci. U.S.A. 86:3389-3393(1989).  
[7]  
RA SEQUENCE OF 195-263 FROM N.A.  
RP MEDLINE=86130593; PubMed=2868718;  
RX Dewji N.N., Wenger D.A., Fujibayashi S., Donoviel M., Esch F.,  
RA Hill F., O'Brien J.S.;  
RT "Molecular cloning of the sphingolipid activator protein-1 (SAP-1),  
RT the sulfate sulfatase activator."  
RL Biochem. Biophys. Res. Commun. 134:989-994(1986).  
[8]  
RA SEQUENCE OF 195-274.  
RP TISSUE=Kidney;  
RX MEDLINE=91006165; PubMed=2209618;  
RA Furst W., Schubert J., Machleidt W., Meyer H.E., Sandhoff K.;  
RT "The complete amino-acid sequences of human ganglioside GM2 activator  
RT protein and cerebroside sulfate activator protein."  
RL Eur. J. Biochem. 192:709-714(1990).  
[9]  
RA SEQUENCE OF 195-274.  
RP MEDLINE=89207118; PubMed=3242555;  
RX Kleinschmidt T., Christomanou H., Braunitzer G.;  
RA "Complete amino-acid sequence of the naturally occurring A2 activator  
RT protein for enzymic sphingomyelin degradation: identity to the  
RT sulfate activator protein (SAP-1)."  
RL Biol. Chem. Hoppe-Seyler 369:1361-1365(1988).  
[10]  
RA SEQUENCE OF 311-390.  
RP MEDLINE=89163077; PubMed=3442600;  
RX Kleinschmidt T., Christomanou H., Braunitzer G.;  
RA "Complete amino-acid sequence and carbohydrate content of the  
RT naturally occurring glucosylceramide activator protein (A1 activator)  
RT absent from a new human Gaucher disease variant."  
RL Biol. Chem. Hoppe-Seyler 368:1571-1578(1987).  
[11]  
RA SEQUENCE OF 407-484.  
RP MEDLINE=89000190; PubMed=3048308;  
RX Furst W., Machleidt W., Sandhoff K.;  
RA "The precursor of sulfate activator protein is processed to three  
RT different proteins."  
RL Biol. Chem. Hoppe-Seyler 369:317-328(1988).  
[12]  
RA PARTIAL SEQUENCE OF 405-484.  
RP MEDLINE=89025876; PubMed=2845979;  
RX Morimoto S., Martin B.M., Kishimoto Y., O'Brien J.S.;  
RA "Saposin D: a sphingomyelinase activator."  
RL Biochem. Biophys. Res. Commun. 156:403-410(1988).  
[13]

RP SEQUENCE OF 17-26.  
RC TISSUE=Milk;  
RX MEDLINE=92068206; PubMed=1958198;  
RA Kondoh K., Hineno T., Sano A., Kakimoto Y.;  
RT "Isolation and characterization of prosaposin from human milk."  
RL Biochem. Biophys. Res. Commun. 181:286-292(1991).  
[14]  
RA PARTIAL SEQUENCE (SAPOSIN B), AND STRUCTURE OF CARBOHYDRATES.  
RP TISSUE=Urine;  
RX MEDLINE=20032116; PubMed=10562467;  
RA Fluharty A.L., Lombardo C., Louis A., Stevens R.L., Whitelegge J.P.,  
Waring A.J., To T., Fluharty C.B., Faull K.F.;  
RT "Preparation of the cerebroside sulfate activator (CSAct or saposin B)  
RT from human urine."  
RL Mol. Genet. Metab. 68:391-403(1999).  
[15]  
RA STRUCTURE OF CARBOHYDRATE ON ASN-215.  
RX MEDLINE=21110404; PubMed=11180632;  
RA Faull K.F., Johnson J., Kim M.J., To T., Whitelegge J.P.,  
Stevens R.L., Fluharty C.B., Fluharty A.L.;  
RT "Structure of the asparagine-linked sugar chains of porcine kidney and  
RT human urine cerebroside sulfate activator protein."  
RL J. Mass Spectrom. 35:1416-1424(2000).  
[16]  
RA SAPOSIN D DISULFIDE BONDS.  
RX MEDLINE=99337688; PubMed=10406958;  
RA Tatti M., Salvioli R., Ciaffoni F., Pucci P., Andolfo A.,  
Amoresano A., Vaccaro A.M.;  
RT "Structural and membrane-binding properties of saposin D."  
RL Eur. J. Biochem. 263:486-494(1999).  
[17]  
RA SAPOSIN B DISULFIDE BONDS.  
RX MEDLINE=22398398; PubMed=12510003;  
RA Ahn V.E., Faull K.F., Whitelegge J.P., Higginson J., Fluharty A.L.,  
Prive G.G.;  
RT "Expression, purification, crystallization, and preliminary X-ray  
RT analysis of recombinant human saposin B."  
RL Protein Expr. Purif. 27:186-193(2003).  
[18]  
RA MASS SPECTROMETRY.  
RC TISSUE=Urine;  
RX MEDLINE=99441404; PubMed=10510427;  
RA Faull K.F., Whitelegge J.P., Higginson J., To T., Johnson J.,  
Krutchinsky A.N., Standing K.G., Waring A.J., Stevens R.L.,  
Fluharty C.B., Fluharty A.L.;  
RT "Cerebroside sulfate activator protein (Saposin B): chromatographic  
RT and electrospray mass spectrometric properties."  
RL J. Mass Spectrom. 34:1040-1054(1999).  
[19]  
RA X-RAY CRYSTALLOGRAPHY (2.2 ANGSTROMS) OF 195-273, AND MUTAGENESIS OF  
RP ILE-240.  
RX MEDLINE=22406333; PubMed=12518053;  
RA Ahn V.E., Faull K.F., Whitelegge J.P., Fluharty A.L., Prive G.G.;  
RT "Crystal structure of saposin B reveals a dimeric shell for lipid  
RT binding."  
RL Proc. Natl. Acad. Sci. U.S.A. 100:38-43(2003).  
[20]  
RA REVIEW ON MLD VARIANTS.  
RP MEDLINE=95170731; PubMed=7866401;  
RX Gieselmann V., Zlotogora J., Harris A., Wenger D.A., Morris C.P.;  
RT "Molecular genetics of metachromatic leukodystrophy."  
RL Hum. Mutat. 4:233-242(1994).  
[21]  
RA VARIANT MLD ILE-217.  
RX MEDLINE=90147748; PubMed=2302219;  
RA Rafi M.A., Zhang X.-L., Degala G., Wenger D.A.;  
RT "Detection of a point mutation in sphingolipid activator protein-1  
RT mRNA in patients with a variant form of metachromatic  
RT leukodystrophy."  
RL Biochem. Biophys. Res. Commun. 166:1017-1023(1990).  
[22]  
RA SEQUENCE FROM N.A., AND VARIANT MLD ILE-217.  
RP MEDLINE=90207231; PubMed=2320574;

```
Query Match 100.0%; Score 2789; DB 1; Length 524;
Best Local Similarity 100.0%; Pred. No. 2e-176;
Matches 524; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MYALFLASLLGALAGVVLGKCTRGSAVWCNQVKTASDCGAVKCLQTVWVKPTVKS 60
Db 1 MYALFLASLLGALAGVVLGKCTRGSAVWCNQVKTASDCGAVKCLQTVWVKPTVKS 60
Qy 61 LPCDICDVTAAAGDMLKDNATEBEILLVLEKTCOMLFPKPNMSASCKEIVDSVLPVILDI 120
Db 61 LPCDICDVTAAAGDMLKDNATEBEILLVLEKTCOMLFPKPNMSASCKEIVDSVLPVILDI 120
Qy 121 IKGEMSRPGEVCSALNLCESLQKHLAELNHQKLESNKIPELDMTEVVPANIPILLY 180
Db 121 IKGEMSRPGEVCSALNLCESLQKHLAELNHQKLESNKIPELDMTEVVPANIPILLY 180
Qy 181 PDGPRSKPQKNGDVQDCIQMTDITQAVRTNSTFVQALVEHVKEECDRLGPGMADI 240
Db 181 PDGPRSKPQKNGDVQDCIQMTDITQAVRTNSTFVQALVEHVKEECDRLGPGMADI 240
Qy 241 CKNIYSQVSEIAIQMHMHQPKETCALVGFCDVEKMPQTLVPAKVASKNVIPEALELVE 300
Db 241 CKNIYSQVSEIAIQMHMHQPKETCALVGFCDVEKMPQTLVPAKVASKNVIPEALELVE 300
Qy 301 PIKKEHPAKSDVYCEVCEFLVKEVTKLIDNNKTEKEILDADFDMCKSKLPKSLSEECQEV 360
Db 301 PIKKEHPAKSDVYCEVCEFLVKEVTKLIDNNKTEKEILDADFDMCKSKLPKSLSEECQEV 360
Qy 361 VDTGSSILSLLEVSPELVCSMLHLCGTRRLPALTVHTVQKGGFCEVCKKLGVYLD 420
Db 361 VDTGSSILSLLEVSPELVCSMLHLCGTRRLPALTVHTVQKGGFCEVCKKLGVYLD 420
Qy 421 RNLEKNTKQBIILAELKGCFLPDPYQKQCDQFVAEVEPVLIILVEVMDPSFVCLKIG 480
Db 421 RNLEKNTKQBIILAELKGCFLPDPYQKQCDQFVAEVEPVLIILVEVMDPSFVCLKIG 480
Qy 481 ACPSAHPLLGTEKICWGPSYWCQNTETAACNAVEHCKRHVWN 524
Db 481 ACPSAHPLLGTEKICWGPSYWCQNTETAACNAVEHCKRHVWN 524

RESULT 2
SAP_BOVIN STANDARD; PRT; 525 AA.
AC P26779; Q9N2G4;
DT 01-AUG-1992 (Rel. 23, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DE Proactivator polypeptide precursor [Contains: Saposin A (Protein A);
DE Saposin B (Sphingolipid activator protein 1) (SAP-1) (Cerebroside
DE sulfate activator) (CSAct) (Dispersin) (Sulfatide/GMI activator);
DE Saposin C (Co-beta-glucosidase) (Al activator) (Glucosylceramidase
DE activator) (Sphingolipid activator protein 2) (SAP-2); Saposin D
DE (Protein C) (Component C)].
GN PSAP.
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovidae; Bovinae; Bos.
OK NCBI_TaxID=9913;
RN [1]
RP SEQUENCE FROM N.A., AND VARIANTS.
RC TISSU=Mammary gland;
RA Azuma N., Yoshida K.;
RT "RT-PCR cloning of bovine prosaposin.";
RL Submitted (JAN-2000) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE OF 312-391.
RC TISSU=Spleen;
RA Sano A., Mizuno T., Kondoh K., Hineno T., Ueno S.-I., Kakimoto Y.,
RA Morita N.;
```

```
RT "Saposin-C from bovine spleen; complete amino acid sequence and
RT relation between the structure and its biological activity.";
RL Biochim. Biophys. Acta 1120:75-80(1992).
CC -!- FUNCTION: The lysosomal degradation of sphingolipids takes place
CC by the sequential action of specific hydrolases. Some of these
CC enzymes require specific low-molecular mass, non-enzymic proteins:
CC the sphingolipids activator proteins (coproteins) (By similarity).
CC -!- FUNCTION: Saposin A and saposin C stimulate the hydrolysis of
CC galactosylceramide by beta-glucosylceramidase (EC 3.2.1.45) and
CC galactosylceramide by beta-galactosylceramidase (EC 3.2.1.46).
CC Saposin C apparently acts by combining with the enzyme and acidic
CC lipid to form an activated complex, rather than by solubilizing
CC the substrate.
CC -!- FUNCTION: Saposin B stimulates the hydrolysis of galacto-
CC cerebroside sulfate by arylsulfatase A (EC 3.1.1.6.8), GMI
CC gangliosides by beta-galactosidase (EC 3.2.1.23) and
CC globotriaosylceramide by alpha-galactosidase A (EC 3.2.1.22).
CC Saposin B forms a solubilizing complex with the substrates of the
CC sphingolipid hydrolases (By similarity).
CC -!- FUNCTION: Saposin D is a specific sphingomyelin phosphodiesterase
CC activator (EC 3.1.4.12) (By similarity).
CC -!- SUBUNIT: Saposin B is a homodimer (By similarity).
CC -!- SUBCELLULAR LOCATION: Lysosomal.
CC -!- PFM: This precursor is proteolytically processed to 4 small
CC peptides, which are similar to each other and are sphingolipid
CC hydrolase activator proteins (By similarity).
CC -!- SIMILARITY: Contains 2 saposin A-type domains.
CC -!- SIMILARITY: Contains 4 saposin B-type domains.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
EMBL; AB036791; BAA95677.1; -.
DR InterPro; IPR003119; SapA.
DR InterPro; IPR007856; SapB 1.
DR InterPro; IPR008138; SapB 2.
DR InterPro; IPR008140; SapB_sub.
DR InterPro; IPR008373; Saposin.
DR InterPro; IPR008139; SaposinB.
DR Pfam; PF02199; SAPA; 2.
DR Pfam; PF05184; SapB 1; 4.
DR Pfam; PF03489; SapB 2; 4.
DR PRINTS; PR01797; SAPOSIN.
DR ProDom; PD001732; SapB_sub; 3.
DR SMART; SM00162; SAPA; 2.
DR SMART; SM00118; SAPB; 4.
KW Signal; Glycoprotein; Lysosome; Sphingolipid metabolism; Repeat.
FT SIGNAL 1 16
FT PROPEP 17 58
FT CHAIN 60 142 SAPOSIN A.
FT PROPEP 144 195
FT CHAIN 196 275 SAPOSIN B.
FT PROPEP 277 310
FT CHAIN 312 392 SAPOSIN C.
FT PROPEP 393 404
FT CHAIN 406 487 SAPOSIN D.
FT PROPEP 489 525
FT DOMAIN 21 54
FT DOMAIN 59 142 SAPOSIN-LIKE TYPE A 1.
FT DOMAIN 194 276 SAPOSIN-LIKE TYPE B 1.
FT DOMAIN 312 393 SAPOSIN-LIKE TYPE B 2.
FT DOMAIN 406 487 SAPOSIN-LIKE TYPE B 3.
FT DOMAIN 492 525 SAPOSIN-LIKE TYPE B 4.
FT DISULFID 63 138 SAPOSIN-LIKE TYPE A 2.
FT DISULFID 66 132 BY SIMILARITY.
FT DISULFID 94 106 BY SIMILARITY.
FT DISULFID 199 272 BY SIMILARITY.
FT DISULFID 202 266 BY SIMILARITY.
```

```

FT DISULFID 231 242 BY SIMILARITY.
FT DISULFID 316 389 BY SIMILARITY.
FT DISULFID 319 383 BY SIMILARITY.
FT DISULFID 347 358 BY SIMILARITY.
FT DISULFID 410 483 BY SIMILARITY.
FT DISULFID 413 477 BY SIMILARITY.
FT DISULFID 441 452 BY SIMILARITY.
FT CARBOHYD 80 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 101 101 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 216 216 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 333 333 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 427 427 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT VARIANT 127 127 H -> R.
FT VARIANT 260 263 MQPK -> IRIR.
FT CONFLICT 317 317 E -> Q (IN REF. 2).
FT CONFLICT 367 367 R -> S (IN REF. 1).
SQ SEQUENCE 525 AA; 58120 MW; 293AFC0FB9C4FA99 CRC64;

Query Match 87.0%; Score 2425.5; DB 1; Length 525;
Best Local Similarity 84.2%; Pred. No. 1.7e-152;
Matches 442; Conservative 45; Mismatches 37; Indels 1; Gaps 1;

Qy 1 MYALFLLASLLGALAGPVLGLKECTRGSAVWQNVKTSADCGAVKHICLOTWNNKPTVKS 60
Db 1 MYSFFVLASLLGALLASPLVGLRECTRGSAVWQNVKTAADCGAVQHCLQTVWNSKPTVKS 60

Qy 61 LPDCICKOVVTAAGDMLKDNATEEELVYLEKTCDWLPKPNMSASCKEIVDSYLPVLDI 120
Db 61 LPDCICKOVITAAGNLLKDNATEQEILMYLERTCDWLKPNMSASCKEIVDSYLPVLDM 120

Qy 121 IKGMSRPGVCSALNLCESLOKHLAELNHQKLESNKIPBLDMTEVVAPFMANIPLLLY 180
Db 121 IKGQSHPGVCSALNLCESLOKHLAELNHQKLESNQIPBLDMAEVVAPFMANIPFLY 180

Qy 181 PDGPRSPQPKD-NGDVQCDCIQMVTDIQTAVRTNSTFVOALVHVHKEECDRLGPGMAD 239
Db 181 PDGSHSPQPKANGVNGVQCDCIQLVTDVQELARTNSTFVEALVDHAKKECDRLGPGMSD 240

Qy 240 ICKNVIQYSEITAIOMVHMHPKEICALVGFCDVEKPMQTLVPAKVASKNVIPAILELV 299
Db 241 MCKNIYINQYSEVIAIQVMVHMHPKEICVLGFCDEKPMQTLVPAEAVSENVIPALGLV 300

Qy 300 EPIKHEVPAKSDVYCEVCEFLVKEVTKLIDNNKTEKEILDAFDKMSCKLPKSLSECOE 359
Db 301 EPIKOPAPAKADIYCEVCEFLVKEVTKLIDNNKTEKEILDAFDKMSCKLPKSLSECOE 360

Qy 360 VVDYTGSSILILEEVSPELVCSMLHLCSGTRLPALTVHVTOFKDGGFCVECKLVGYL 419
Db 361 VVDYTGSSILILEEVSPELVCSMLHLCSGTRLPALTVHVTOFKDGGFCVECKLVGYL 420

Qy 420 DRNLEKSTKQEIILAALEKGSFLPDVPYKOCQOFVAEYEPVLIETILVEWDPFSVCLKI 479
Db 421 DRNLEKSTKQEIILAALEKGSFLPDYQKQCOQFVTEYEPVLIETILVEWDPFSVCLKI 480

Qy 480 GACSAHKPLGTBKCITWSPGYWQNTTETAQCNNAVEHCKRHHVN 524
Db 481 GACPAHKPLTGAEEKVWSPGYWQNMESAAALCNNAVEHCHRRHVN 525

RESULT 3
SAP_RAT
ID SAP_RAT STANDARD; PRT; 554 AA.
AC P10560; 062841; 064190;
DT 01-JUL-1989 (Rel. 11, Created)
DT 01-JUL-1989 (Rel. 11, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE Sulfated glycoprotein 1 precursor (SGP-1) (Prosaposin).
GN PSAP OR SGP1
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]

```

```

RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
RC TISSUE=Sertoli cells;
RX MEDLINE=85000647; PubMed=3048385;
RA Collard M.W., Sylvester S.R., Tsuruta J.K., Griswold M.D.;
RT "Biosynthesis and molecular cloning of sulfated glycoprotein 1
RT secreted by rat Sertoli cells: sequence similarity with the
RT 70-kilodalton precursor to sulfatide/GM1 activator.";
RL Biochemistry 27:4557-4564(1988).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Testis;
RX MEDLINE=96128541; PubMed=8573994;
RA Morales C.R., El-Alfy M., Zhao Q., Igdowna S.A.;
RT "Molecular role of sulfated glycoprotein-1 (SGP-1/prosaposin) in
RT Sertoli cells.";
RL Histol. Histopathol. 10:1023-1034(1995).
RN [3]
RP SEQUENCE FROM N.A.
RC TISSUE=Testicle;
RX MEDLINE=96175245; PubMed=8601692;
RA Morales C.R., El-Alfy M., Zhao Q., Igdowna S.A.;
RT "Expression and tissue distribution of rat sulfated glycoprotein-1
RT (prosaposin).";
RL J. Histochem. Cytochem. 44:327-337(1996).
CC -1- SUBCELLULAR LOCATION: Extracellular.
CC -1- SIMILARITY: Contains 2 saposin A-type domains.
CC -1- SIMILARITY: Contains 4 saposin B-type domains.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch)
CC -----
DR EMBL; M19936; AAA42136.1; -
DR EMBL; S81353; AAB36042.2; -
DR EMBL; S81373; AAB36233.2; -
DR PIR; A28716; A28716.
DR InterPro; IPR003119; Sapa.
DR InterPro; IPR007856; SapaB.
DR InterPro; IPR008138; SapaB.2.
DR InterPro; IPR008140; SapaB.sub.
DR InterPro; IPR008373; Saposin.
DR InterPro; IPR008139; SaposinB.
DR Pfam; PF02199; Sapa.2.
DR Pfam; PF05184; SapaB.1; 3.
DR Pfam; PF03489; SapaB.2; 4.
DR PRINTS; PR01797; SAPOSIN.
DR ProDom; PD001732; SapaB.sub.3.
DR SMART; SM00162; SAPA.2.
DR SMART; SM00118; SAPA.4.
DR Signal; Glycoprotein; Repeat.
FT SIGNAL 1 16
FT CHAIN 17 554 SULFATED GLYCOPROTEIN 1.
FT DOMAIN 21 54 SAPOSIN-LIKE TYPE A 1.
FT DOMAIN 61 138 SAPOSIN-LIKE TYPE B 1.
FT DOMAIN 193 274 SAPOSIN-LIKE TYPE B 2.
FT DOMAIN 310 391 SAPOSIN-LIKE TYPE B 3.
FT DOMAIN 435 516 SAPOSIN-LIKE TYPE B 4.
FT DOMAIN 521 554 SAPOSIN-LIKE TYPE A 2.
FT DISULFID 63 138 BY SIMILARITY.
FT DISULFID 66 132 BY SIMILARITY.
FT DISULFID 94 106 BY SIMILARITY.
FT DISULFID 197 270 BY SIMILARITY.
FT DISULFID 200 264 BY SIMILARITY.
FT DISULFID 229 240 BY SIMILARITY.
FT DISULFID 314 387 BY SIMILARITY.
FT DISULFID 317 381 BY SIMILARITY.
FT DISULFID 345 356 BY SIMILARITY.
FT DISULFID 439 512 BY SIMILARITY.
FT DISULFID 442 506 BY SIMILARITY.

```





```
FT DISULFID 473 484 BY SIMILARITY.
FT CARBOHYD 80
FT CARBOHYD 214
FT CARBOHYD 334
FT CARBOHYD 459
FT CARBOHYD 83
FT CARBOHYD 158
FT CARBOHYD 160
FT CARBOHYD 171
FT CARBOHYD 244
FT CARBOHYD 254
FT CARBOHYD 255
FT CARBOHYD 260
FT CARBOHYD 307
FT CARBOHYD 322
FT CARBOHYD 349
FT CARBOHYD 367
FT CARBOHYD 370
FT CARBOHYD 373
FT CARBOHYD 391
FT CARBOHYD 393
FT CARBOHYD 406
FT CARBOHYD 430
FT CARBOHYD 445
FT CARBOHYD 448
FT CARBOHYD 557
SQ SEQUENCE 557 AA; 61422 MW; 134593E20499E35E CRC64;

Query Match
Best Local Similarity 69.5%; Score 1938.5; DB 1; Length 557;
Matches 355; Conservative 78; Mismatches 90; Indels 35; Gaps 3;

QY 1 MYALFLLSLLGAAAGPVLGKECTRGSAVYQVKTASDCGAVGHGCLQVWKNKPTVKS 60
Db 1 MYALFLLSLLGAAAGPVLGKECTRGSAVYQVKTASDCGAVGHGCLQVWKNKPTVKS 60
QY 61 LPEDICKVWTAAGDMLKDNATEREILVYLEKTCOWLPKPNNSASCKEIVDSYLPVLDI 120
Db 61 LPEDICKVWTAAGDMLKDNATEREILVYLEKTCOWLPKPNNSASCKEIVDSYLPVLDI 120
QY 121 IGENSRGFEVCSALNLCESLQHLAEHLNQHLESNKIPBLDMTEVVAPFMANIPALLY 180
Db 121 IGENSRGFEVCSALNLCESLQHLAEHLNQHLESNKIPBLDMTEVVAPFMANIPALLY 180
QY 181 PDGPRSPKPKDGVDCODCIQWTDIQTAVRTNFTVQALVEHVKEECDRLPGMGADI 240
Db 181 PDGPRSPKPKDGVDCODCIQWTDIQTAVRTNFTVQALVEHVKEECDRLPGMGADI 240
QY 241 CKNYISQYSEIAIQMMHMM---QPKCALVGFCDVEKEMPMOTLVPKASKNVTPALE 297
Db 241 CKNYISQYSEIAIQMMHMM---QPKCALVGFCDVEKEMPMOTLVPKASKNVTPALE 297
QY 298 LVEPIKKEHVPAKSDVYCEVEFLVKEVTKLIDNNKTEKELIDAFKMCCKLPKSLSEBC 357
Db 298 LVEPIKKEHVPAKSDVYCEVEFLVKEVTKLIDNNKTEKELIDAFKMCCKLPKSLSEBC 357
QY 358 QEVVDYTGSSILSLLEVEVPELVCSMLHLCSG----- 390
Db 358 QEVVDYTGSSILSLLEVEVPELVCSMLHLCSG----- 390
QY 360 QEVVGTGSPSLDIFIEHVPSSLCVGLCAARPELVEALEOPAPAIVSALLKEPTPK 419
Db 360 QEVVGTGSPSLDIFIEHVPSSLCVGLCAARPELVEALEOPAPAIVSALLKEPTPK 419
QY 391 ----TRLPALTVHTVQPKDGGCEVCKVGLYLDRLNLEKNSTKQELIAALEKGCFLPDP 446
Db 391 ----TRLPALTVHTVQPKDGGCEVCKVGLYLDRLNLEKNSTKQELIAALEKGCFLPDP 446
QY 447 YQKQCDQFVAYEPVLEIIVVMDSPFVCLKIGACPSAKHPLLTGTEKINGSPSWQONT 506
Db 447 YQKQCDQFVAYEPVLEIIVVMDSPFVCLKIGACPSAKHPLLTGTEKINGSPSWQONT 506
QY 480 YQKQCDQFVAYEPVLEIIVVMDSPFVCSKIGVCPKSAKLLGTEKCVGSPWYSCQNM 539
Db 480 YQKQCDQFVAYEPVLEIIVVMDSPFVCSKIGVCPKSAKLLGTEKCVGSPWYSCQNM 539
QY 507 ETRAAQCNVABCHRRHVN 524
Db 507 ETRAAQCNVABCHRRHVN 524
QY 540 ETRAAQCNVABCHRRHVN 557
Db 540 ETRAAQCNVABCHRRHVN 557
```

```
SAP_CHICK
ID SAP_CHICK STANDARD; PRT; 518 AA.
AC O13035;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Proactivator polypeptide precursor [Contains: Saposin A; Saposin B;
DE Saposin C; Saposin D].
DE PSAP.
GN Gallus gallus (Chicken).
OS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC Gallus.
OX NCBI_TaxID=9031;
RN [1]
RP SEQUENCE FROM N.A., AND SEQUENCE OF 194-203.
RC TISSUE=Brain, and Liver;
RX MEDLINE=98129745; PubMed=9461526;
RA Azuma N., Seo H.-C., Lie O., Fu Q., Gould R.M., Hiraiwa M., Burt D.W.,
RA Paton I.R., Morrison D.R., O'Brien J.S., Kishimoto Y.;
RA "Cloning, expression and map assignment of chicken prosaposin.";
RL Biochem. J. 330:321-327(1998).
RN [2]
RP SEQUENCE FROM N.A.
RA Altman N., Horowitz M.;
RL Submitted (NOV-1998) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: The lysosomal degradation of sphingolipids takes place
CC by the sequential action of specific hydrolases. Some of these
CC enzymes require specific low-molecular mass, non-enzymic proteins:
CC the sphingolipids activator proteins (coproteins) (By similarity).
CC -!- FUNCTION: Saposin A and saposin C stimulate the hydrolysis of
CC glucosylceramide by beta-glucosylceramidase (EC 3.2.1.45) and
CC saposin C apparently acts by combining with the enzyme and acidic
CC lipid to form an activated complex, rather than by solubilizing
CC the substrate (By similarity).
CC -!- FUNCTION: Saposin B stimulates the hydrolysis of galacto-
CC cerebroside sulfate by arylsulfatase A (EC 3.1.6.8), GM1
CC gangliosides by beta-galactosidase (EC 3.2.1.23) and
CC glioblastosylceramide by alpha-galactosidase A (EC 3.2.1.22).
CC Saposin B forms a solubilizing complex with the substrates of the
CC sphingolipid hydrolases (By similarity).
CC -!- FUNCTION: Saposin D is a specific sphingomyelin phosphodiesterase
CC activator (EC 3.1.4.12) (By similarity).
CC -!- SUBUNIT: Saposin B is a homodimer (By similarity).
CC -!- SUBCELLULAR LOCATION: Lysosomal (By similarity).
CC -!- PTM: This precursor is proteolytically processed to 4 small
CC peptides, which are similar to each other and are sphingolipid
CC hydrolase activator proteins (By similarity).
CC -!- SIMILARITY: Contains 2 saposin A-type domains.
CC -!- SIMILARITY: Contains 4 saposin B-type domains.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC -----
DR EMBL; AB003471; BAA19914.1; -.
DR EMBL; AF108656; AAF05899.1; -.
DR InterPro; IPR003119; Sapa.
DR InterPro; IPR007856; SapaB_1.
DR InterPro; IPR008138; SapaB_2.
DR InterPro; IPR008140; SapaB_sub.
DR InterPro; IPR008375; Saposin.
DR InterPro; IPR008139; SaposinB.
DR Pfam; PF02199; Sapa; 2.
DR Pfam; PF05184; SapaB_1; 4.
DR Pfam; PF03489; SapaB_2; 4.
DR PRINTS; PR01797; Saposin.
DR ProDom; PD001732; SapaB_sub; 3.
```

KW Signal; Glycoprotein; Lysosome; Sphingolipid metabolism; Repeat;  
KW GM2-gangliosidosis.

```
FT SIGNAL 1 17 POTENTIAL.
FT PROPEP 18 60
FT CHAIN 61 143 SAPOSIN A.
FT PROPEP 145 193
FT CHAIN 194 276 SAPOSIN B.
FT PROPEP 278 305 SAPOSIN C.
FT CHAIN 307 387 SAPOSIN D.
FT PROPEP 389 398
FT CHAIN 399 480
FT PROPEP 482 518
FT DOMAIN 22 55 SAPOSIN-LIKE TYPE A 1.
FT DOMAIN 60 143 SAPOSIN-LIKE TYPE B 1.
FT DOMAIN 193 277 SAPOSIN-LIKE TYPE B 2.
FT DOMAIN 307 388 SAPOSIN-LIKE TYPE B 3.
FT DOMAIN 399 480 SAPOSIN-LIKE TYPE B 4.
FT DOMAIN 485 518 SAPOSIN-LIKE TYPE A 2.
FT DISULFID 64 139 BY SIMILARITY.
FT DISULFID 67 133 BY SIMILARITY.
FT DISULFID 95 107 BY SIMILARITY.
FT DISULFID 197 273 BY SIMILARITY.
FT DISULFID 200 267 BY SIMILARITY.
FT DISULFID 229 240 BY SIMILARITY.
FT DISULFID 311 384 BY SIMILARITY.
FT DISULFID 314 378 BY SIMILARITY.
FT DISULFID 342 353 BY SIMILARITY.
FT DISULFID 403 476 BY SIMILARITY.
FT DISULFID 406 470 BY SIMILARITY.
FT DISULFID 434 445 BY SIMILARITY.
FT CARBOHYD 81 81 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 214 214 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 328 328 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 420 420 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CONFLICT 94 94 R -> T (IN REF. 2).
FT CONFLICT 486 486 E -> D (IN REF. 2).
SQ SEQUENCE 518 AA; 57601 MW; B8030008891C3963 CRC64;
```

```
Query Match 60.2%; Score 1679.5; DB 1; Length 518;
Best Local Similarity 57.4%; Pred. No. 2.2e-103;
Matches 301; Conservative 96; Mismatches 114; Indels 13; Gaps 5;

Qy 4 LFLASLLGALAGVILGKECTRGSAVWCQNVKTASDCGAVKHCLQTVNKNKPTVKS LPC 63
Db 5 LTLTLLGALAAVSPVLWQDCAKGPEVWCQSLRTASQCGAVKHCCQNVKSPAVNSIPC 64

Qy 64 DICKDVTAGDMLKDNATBEEILVLEKTCMDLPKPNMGSASCKEIVDSYLPVLLDIKG 123
Db 65 DLCKELVTVVGKVLKDRGTEDTIRSYLEKRCFPLDQGLASECKEIVDSYLPVIMDIKE 124

Qy 124 EMSRPGVCSALNLCESLQKHLAELNHOKLESNKIPELDMTEVVVAFPMANIPLLLYPOD 183
Db 125 EFDKPEVVCALSQCQSLQKHLAAMKLOKQLOSINKIPELDFSELTSPFMANVPLLLYPQD 184

Qy 184 GPRSKPQKONGVQDCPCIQWVTDIQTAVRTNSTFVQALVEHVKKECDRLGPGMADICKN 243
Db 185 KPKQK--SKATEDVCQDCIRLVTVDQAVRTNAFVKSVAHAKEECDRLGPGMSDWCKS 242

Qy 244 YISQYSIAIQMMHMM--QPKETCALVGCDFEVENPMQTLVPAKVASKNVIPAELVE 300
Db 243 YISEYSLAIQMMHMKDQDPKDCANVGFCSVKSVPLOTLVPAQVHHE-----VKME 296

Qy 301 PIKHEVPKASDVYCEVCEFLAVKEVTKLIDNNKTEKILDAFDKWCCKLPKLSLSEECQEV 360
Db 297 TVEKATVQEKTFVCEICETWVKRVTGLESNKTEETEEIVHEMVEVVCYLLPASVKDQCKDF 356

Qy 361 VDTYGSISILLBEVSPVCSMLHLCSGTRLPALTAVHYTPQKDGGEVCEVKLGYLD 420
Db 357 IEVYQALIDMLLEATNPVAVCMVKCAANKPQPPV-VVKPA-GGFCDCIKMIVAYAD 414

Qy 421 RNLEKNTKQIILAALEKGSFLDPDPQKQCDQFVAEYEPVLEILVEVMDPSPVCLKIG 480
Db 415 KELEKNATTTIEALLEKVFHLEPSVSDQCVQVEYEPVWVQVLLAEMMDPTFVCTKLG 474
```

```
Qy 481 ACPSAHKPLLGTEKCIWGPSYWCQNTETATAQCNAVEHCKRHVWN 524
Db 475 VCGAANKPLLGEDACVWGPYCKWCKMWTAAQCNAVDHCRHVWN 518

RESULT 6
SAP_PIG ID SAP_PIG STANDARD; PRT; 80 AA.
AC P81405;
DT 15-DEC-1998 (Rel. 37, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Saposin B (Cerebroside sulfate activator) (CS-ACT) (Non-specific
DE activator) (Sphingolipid activator protein 1) (SAP-1).
OS Sus scrofa (Pig).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
OX NCBI_TaxID=9823;
RN [1]
RC TISSUE=Kidney;
RC TISSUE=Kidney;
RX MEDLINE=93229506; PubMed=8471613;
RA Stevens R.L., Faull K.F., Conklin K.A., Green B.N., Fluharty A.L.;
RT "Porcine cerebroside sulfate activator: further structural
RT characterization and disulfide identification.";
RL Biochemistry 32:4051-4059(1993).
RN [2]
RP SEQUENCE OF 1-64.
RC TISSUE=Kidney;
RX MEDLINE=92222651; PubMed=1562358;
RA Fluharty A.L., Katona Z., Meek W.E., Frei K., Fowler A.V.;
RT "The cerebroside sulfate activator from pig kidney: purification and
RT molecular structure.";
RL Biochem. Med. Metab. Biol. 47:66-85(1992).
RN [3]
RP STRUCTURE OF CARBOHYDRATE ON ASN-21.
RX MEDLINE=21110404; PubMed=11180632;
RA Faull K.F., Johnson J., Kim M.J., To T., Whitelegge J.P.,
RA Krutchinsky A.N., Standing K.G., Waring A.J., Stevens R.L.,
RA Fluharty C.B., Fluharty A.L.;
RT "Cerebroside sulfate activator protein (Saposin B): chromatographic
RT and electrospray mass spectrometric properties.";
RL J. Mass Spectrom. 34:1040-1054(1999).
CC -1- FUNCTION: Saposin B stimulates the hydrolysis of galacto-
CC cerebroside sulfate by arylsulfatase A (EC 3.1.6.8), GM1
CC gangliosides by beta-galactosidase (EC 3.2.1.23) and
CC globotriaosylceramide by alpha-galactosidase A (EC 3.2.1.22).
CC Saposin B forms a solubilizing complex with the substrates of the
CC sphingolipid hydrolases.
CC -1- SUBUNIT: Saposin B is a homodimer (By similarity).
CC -1- PTM: The one residue extended Saposin B-Val is only found in a
CC minority of the chains.
CC -1- SIMILARITY: Contains 1 saposin B-type domain.
DR GlycoSuiteDB; P81405;
DR InterPro; IPR007856; SapB_1.
DR InterPro; IPR008138; SapB_2.
DR InterPro; IPR008373; Saposin.
DR InterPro; IPR008139; SaposinB.
DR Pfam; PF05184; SapB_1; 1.
DR Pfam; PF03489; SapB_2; 1.
DR PRINTS; PR01797; SAPOSIN.
DR SMART; SM00118; SAPB; 1.
KW Glycoprotein; Sphingolipid metabolism.
```

```
FT CHAIN 1 79 SAPOSIN B.
FT CHAIN 1 80 SAPOSIN B-VAL.
FT DOMAIN 1 80 SAPOSIN-LIKE TYPE B.
FT DISULFID 4 77
FT DISULFID 7 71
FT DISULFID 36 47
FT CARBOHYD 21 21
SQ SEQUENCE 80 AA; 8949 MW; EFTBA49B63E789C CRC64;

Query Match 14.1%; Score 393; DB 1; Length 80;
Best Local Similarity 88.8%; Pred. No. 1.1e-19;
Matches 71; Conservative 5; Mismatches 4; Indels 0; Gaps 0;

QY 195 GDVQCDCIQMVTDTQTAVRTNSTFVQALVHVKECRDLRGLPGMADICKNYSQYSEIAIQ 254
DB 1 GDVQCDCIQMVTDTQAVRTNSTFVQALVHVKECRDLRGLPGMADICKNYSQYSEIAIQ 60

QY 255 MMHMQPKDICALVGFCEV 274
DB 61 MMHMQPKDICALVGFCEV 80

RESULT 7
PSPB_MOUSE STANDARD; PRT; 377 AA.
AC P50405;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Pulmonary surfactant-associated protein B precursor (SP-B) (Pulmonary surfactant-associated proteolipid SPL(Phe)).
DE SFTPB OR SFTPB.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]_TaxID=10090;
RP SEQUENCE FROM N.A.
RC STRAIN=DBA/2J; TISSUE=Liver;
RX MEDLINE=95208782; PubMed=7900819;
RA Bruno M.A., Bohinski R.J., Carter J.E., Foss K.A., Whitsett J.A.;
RT "Structure and function of the mouse surfactant protein B gene.";
RL Am. J. Physiol. 268:1381-1389(1995)
CC -!- FUNCTION: PULMONARY SURFACTANT-ASSOCIATED PROTEINS PROMOTE ALVEOLAR STABILITY BY LOWERING THE SURFACE TENSION AT THE AIR-LIQUID INTERFACE IN THE PERIPHERAL AIR SPACES. SP-B INCREASES THE COLLAPSE PRESSURE OF PALMITIC ACID TO NEARLY 70 MILLINEWTONS PER METER.
CC -!- SUBUNIT: Homodimer; disulfide-linked.
CC -!- SUBCELLULAR LOCATION: Extracellular.
CC -!- MISCELLANEOUS: Pulmonary surfactant consists of 90% lipid and 10% protein. There are 4 surfactant-associated proteins: 2 collagenous, carbohydrate-binding glycoproteins (SP-A and SP-D) and 2 small hydrophobic proteins (SP-B and SP-C).
CC -!- SIMILARITY: Contains 1 saposin A-type domain.
CC -!- SIMILARITY: Contains 3 saposin B-type domains.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).
CC -----
DR EMBL; S78114; AAB34846.2; -
DR HSP; P07988; IDFW.
DR MGD; MGI:109516; Sftpb.
DR InterPro; IPR003119; Sapa.
DR InterPro; IPR007856; SApB.1.
DR InterPro; IPR008138; SApB.2.
DR InterPro; IPR008140; SApB_sub.
DR -----
```

```
DR InterPro; IPR008373; Saposin.
DR InterPro; IPR008139; SaposinB.
DR InterPro; IPR008137; Surfactant_B.
DR Pfam; PF02199; SAPA; 1.
DR Pfam; PF05184; SApB_1; 1.
DR Pfam; PF03489; SApB_2; 3.
DR PRINTS; PR01797; SAPOSIN.
DR ProDom; PD001732; SApB_sub; 1.
DR ProDom; PD008002; Surfactant_B; 1.
DR SMART; SM00162; SAPA; 1.
DR SMART; SM00118; SApB; 3.
KW Surface film; Gaseous exchange; Glycoprotein; Repeat.
FT PROPEP 1 191 PULMONARY SURFACTANT-ASSOCIATED PROTEIN
FT CHAIN 192 270 B.
FT PROPEP 271 377
FT DOMAIN 27 60 SAPOSIN-LIKE TYPE A.
FT DISULFID 199 268 BY SIMILARITY.
FT DISULFID 202 262 BY SIMILARITY.
FT DISULFID 226 237 BY SIMILARITY.
FT DISULFID 239 239 INTERCHAIN (BY SIMILARITY).
FT CARBOHYD 307 307 N-LINKED (GLCNAC...) (POTENTIAL).
SQ SEQUENCE 377 AA; 41728 MW; CB687A82BA3FC56C CRC64;

Query Match 13.7%; Score 382; DB 1; Length 377;
Best Local Similarity 22.9%; Pred. No. 4e-18;
Matches 118; Conservative 75; Mismatches 154; Indels 168; Gaps 18;

QY 4 LFLASLL--GAALAGPVLGLKECTRGSAVWQVKTASDCGAVKHCLQTVNKKTVKSL 61
DB 10 LLLLPFLCCGGAITS--ASSLECAQPGQFWQCSLEHAHVQCRALGCLQEVGHAGANDL 67

QY 62 PCDCIKDVVTAAGDMLKDNATEEILVYLEKTCWLPKPNMSASCKEIVDSVLPVLDII 121
DB 68 -QCEEDI VHLTKTKEDAFQEARKELEQECDDLPLKLVPRCQVDVLPVLDIYF 126

QY 122 KGEMSRPGEVCSALNLCESLOKHLAELNHQKLESNKIPELDMTEVAVPMANIPLLLYP 181
DB 127 QSQIN-PKAINHVGLC----- 142

QY 182 QDGRSKPQKNDGVDVQCDCIQMVTDTQTAVRTNSTFVQALVHVKECRDLRGLPGMADIC 241
DB 143 ---PRGAQKPN-----PGMPD-- 157

QY 242 KNYISQYSEIALQMMHMQPKDICALVGFCEVDEKMPMTQTLVPAKVASKNVIPALE--L 298
DB 158 -----AVPNPLDLKLVLPVPGALL 177

QY 299 VEP-----IKKHEVPAKSDVYCEVCEFLVKEVTKLIDNNKTEKILD-AFDKMKCKLPK 351
DB 178 ARPGHTQDFSEQLPIPLP-FCWLCTRLTKRQVAVI-----PKGVLAVAVSQVCHVVP 231

QY 352 SLSECEQEVVDYTGSSILSILLESPELVCSMLHLC-----GTRLPALTVVHTQ-PKD 405
DB 232 VVGICQCCLABRYTVLLDALLGRVWPQLVGLVLRCTSDAMGPAVPEPLIEEWPLO 291

QY 406 GGFCEVCKLVGLDRNLKNSKTKQELAALEKGC-SFLPDPYQKQCDQVAYEVPVLE 464
DB 292 DTECHFCRSVI-----NQAWNTSEQAMPOAMHQAQLRFLWD--RQKEQFVEQHMPQLLA 344

QY 465 ILVEWMDPSFVCLKIGACPSAHKPLLLGTEKCIWGP 499
DB 345 LVPRSDAHITCQALGVCEAPASPL-----QCFTP 375

RESULT 8
PSPB_MOUSE STANDARD; PRT; 376 AA.
ID PSPB_MOUSE
AC P22355;
DT 01-AUG-1991 (Rel. 19, Created)
DT 01-AUG-1991 (Rel. 19, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Pulmonary surfactant-associated protein B precursor (SP-B) (Pulmonary
```



RC TISSUE=Brain; PubMed=12477932;  
RX MEDLINE=22388257; PubMed=12477932;  
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,  
RA Klausner R.D., Collins F.S., Wagner L., Shennan C.M., Schuler G.D.,  
RA Altschul S.F., Zeeberg H., Moore T., Max S.I., Wang J., Hsieh F.,  
RA Hopkins R.F., Jordan K., Marusina K., Farmer A.A., Rubin G.M., Hong L.,  
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,  
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Schetz T.E.,  
RA Brownstein M.J., Ustin T.B., Tohiyuki S., Carninci P., Frange C.,  
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullany S.J.,  
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,  
RA Villalon D.K., Muzny K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,  
RA Richards S., Worley K.C., Sodergren E.J., Lu X., Gibbs R.A.,  
RA Fahy J., Helton E., Kettelman M., Madan A., Rodrigues S., Sanchez A.,  
RA Blakesley R.W., Touchman A.C., Shevchenko Y., Bouffard G.G.,  
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,  
RA Butterfield Y.S.N., Krzywicki M.I., Skalska U., Smalil D.E.,  
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.,  
RT "Generation and initial analysis of more than 15,000 full-length  
human and mouse cDNA sequences.";  
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).  
RN [5]  
RP SEQUENCE OF 99-381 FROM N.A.  
RX MEDLINE=87231940; PubMed=3035561;  
RA Glasser S.W., Korfhagen T.R., Weaver T., Pilot-Matias T., Fox J.L.,  
RA Whitsett J.A.;  
RT "cDNA and deduced amino acid sequence of human pulmonary surfactant-  
associated proteolipid SPL(Phc).";  
RL Proc. Natl. Acad. Sci. U.S.A. 84:4007-4011 (1987).  
RN [6]  
RP SEQUENCE OF 139-381 FROM N.A.  
RX MEDLINE=88139786; PubMed=3343343;  
RA Revak S.D., Merritt T.A., Degryse E., Stefani L., Courtney M.,  
RA Hallman M., Cochran C.G.;  
RT "Use of human surfactant low molecular weight apoproteins in the  
reconstitution of surfactant biologic activity.";  
RL J. Clin. Invest. 81:826-833 (1988).  
RN [7]  
RP SEQUENCE OF 201-279, AND DISULFIDE BONDS.  
RX MEDLINE=92233937; PubMed=1568474;  
RA Johansson J., Joernvall H., Curstedt T.,  
RT "Human surfactant polypeptide SP-B. Disulfide bridges, C-terminal  
end, and peptide analysis of the airway form.";  
RL FEBS Lett. 301:165-167 (1992).  
RN [8]  
RP STRUCTURE BY FTIR OF 201-225.  
RX MEDLINE=20256237; PubMed=10798379;  
RA Gordon L.M., Lee K.Y., Lipp M.M., Zasadzinski J.A., Walther F.J.,  
RA Sherman M.A., Waring A.J.;  
RT "Conformational mapping of the N-terminal segment of surfactant  
protein B in lipid using 13C-enhanced Fourier transform infrared  
spectroscopy.";  
RL J. Pept. Res. 55:330-347 (2000).  
RN [9]  
RP VARIANT ILE-131.  
RX MEDLINE=20525299; PubMed=11076040;  
RA Lin Z., Pearson C., Chinchilli V., Pietschmann S.M., Luo J., Pison U.,  
RA Floros J.;  
RT "Polymorphisms of human SP-A, SP-B, and SP-D genes: association of  
SP-B Thr311le with ARDS.";  
RL Clin. Genet. 58:181-191 (2000).  
CC -1- FUNCTION: PULMONARY SURFACTANT-ASSOCIATED PROTEINS PROMOTE  
ALVEOLAR STABILITY BY LOWERING THE SURFACE TENSION AT THE AIR-  
LIQUID INTERFACE IN THE PERIPHERAL AIR SPACES. SP-B INCREASES  
THE COLLAPSE PRESSURE OF PALMITIC ACID TO NEARLY 70 MILLINEWTONS  
PER METER.  
CC -1- SUBUNIT: Homodimer; disulfide-linked.  
CC -1- SUBCELLULAR LOCATION: Extracellular.  
CC -1- POLYMORPHISM: There seems to be an association between the Ile-131  
polymorphism and acute respiratory distress syndrome (ARDS).  
CC -1- DISEASE: Defects in SFTPB are a cause of congenital pulmonary  
alveolar proteinosis (PAP) [MIN:265120]. PAP is an autosomal  
recessive fatal respiratory disease.  
CC MISCELLANEOUS: Pulmonary surfactant consists of 90% lipid and 10%  
protein. There are 4 surfactant-associated proteins: 2 collagenous,  
carbohydrate-binding glycoproteins (SP-A and SP-D) and 2 small  
hydrophobic proteins (SP-B and SP-C).  
CC -1- SIMILARITY: Contains 1 saposin A-type domain.  
CC -1- SIMILARITY: Contains 3 saposin B-type domains.  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
between the Swiss Institute of Bioinformatics and the EMBL outstation -  
the European Bioinformatics Institute. There are no restrictions on its  
use by non-profit institutions as long as its content is in no way  
modified and this statement is not removed. Usage by and for commercial  
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC EMBL; M24461; AAB59541.1; -  
DR EMBL; J02761; AAA60212.1; -  
DR EMBL; AF400074; AAK77913.1; -  
DR EMBL; BC032785; AAH32785.1; -  
DR EMBL; M16764; AAA88099.1; ALT\_INIT.  
DR EMBL; M19097; AAA36628.1; -  
DR PIR; A31361; LNHUB  
DR PDB; 1DFW; 24-OCT-01.  
DR PDB; 1KWE; 07-AUG-02.  
DR Genew; HGNC:10801; SFTPB.  
DR MIM; 178640; -  
DR MIM; 265120; -  
DR GO; GO:0007397; P:histogenesis and organogenesis; TAS.  
DR GO; GO:0007585; P:respiratory gaseous exchange; TAS.  
DR InterPro; IPR003119; SAPA\_1.  
DR InterPro; IPR007856; SAPB\_1.  
DR InterPro; IPR008138; SAPB\_2.  
DR InterPro; IPR008140; SAPB\_sub.  
DR InterPro; IPR008373; Saposin.  
DR InterPro; IPR008139; SaposinB.  
DR InterPro; IPR008137; Surfactant\_B.  
DR Pfam; PF02199; SAPA; 1.  
DR Pfam; PF05184; SAPB\_1; 1.  
DR Pfam; PF03489; SAPB\_2; 3.  
DR PRINTS; PD01797; SAPOSIN.  
DR ProDom; PD001732; SapB\_sub; 1.  
DR ProDom; PD008002; Surfactant\_B; 1.  
DR SMART; SM00162; SAPA; 1.  
DR SMART; SM00118; SAPB; 3.  
KW Surface film; Gaseous exchange; Glycoprotein; Repeat; Polymorphism;  
KW 3D-structure.  
FT PROPEP 1 200  
FT CHAIN 201 279  
FT PULMONARY SURFACTANT-ASSOCIATED PROTEIN  
FT B.  
FT PROPEP 280 381  
FT DOMAIN 28 61  
FT DISULFID 208 277  
FT DISULFID 211 271  
FT DISULFID 235 246  
FT DISULFID 248 248  
FT CARBOHYD 129 129  
FT CARBOHYD 311 311  
FT VARIANT 131 131  
FT VARIANT 176 176  
FT VARIANT 228 228  
FT VARIANT 228 228  
FT VARIANT 272 272  
FT CONFLICT 178 178  
FT CONFLICT 318 318  
FT HELIX 208 221  
FT SEQUENCE 381 AA; 42117 MW; 9FD7F66678A35153 CRC64;  
SAPOSIN-LIKE TYPE A.  
INTERCHAIN.  
N-LINKED (GLCNAC. . .) (POTENTIAL).  
N-LINKED (GLCNAC. . .) (POTENTIAL).  
T -> I.  
/FTID=VAR\_006948.  
L -> F.  
/FTID=VAR\_013099.  
A -> R.  
/FTID=VAR\_006949.  
A -> I.  
/FTID=VAR\_006950.  
R -> H.  
/FTID=VAR\_013100.  
L -> V (IN REF. 6).  
P -> L (IN REF. 5).  
9FD7F66678A35153 CRC64;

```
Query Match      12.5%; Score 349.5; DB 1; Length 381;
Best Local Similarity 21.6%; Pred. No. 5.6e-16;
Matches 111; Conservative 72; Mismatches 171; Indels 161; Gaps 16;

QY 2 YALFLLASLGAALAGPVGLKECTRGSAVWQNVKTSADCGAVKHCLQTVWKNKPTVKS 61
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 9 WLLLLLPTLCGPTAAWTTSSLAQGGPFWCQSLQALQCRALGHCLQEVGHWGADDL 68
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 62 PCDICKDVVTAAGDMLKDNATEEILVYLEKTDWLPKPNMSASCKEIVDSYLPVILDI 121
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 69 -CQCECDIVHILNKWAKEAIFQDTRKFELEQECNVLPKLKALMPQCNQVLDYFFLVIDYF 127
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 122 KGEISRPGVCSALNLCESLQKHLAEFLNKHQKLESNKIPELDMTEVAVPFWANIPLLLYP 181
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 128 QNQTDSNG-ICMHLGLCKSRQ----- 147
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 182 QDGRPSRKPQKNGDVCQDCIQMTDITQAVRTNSTFTVQALVEHVKECDRLGPGMADIC 241
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 148 -----PEPQEQ----- 158
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 242 KNYISQYSEIAIQMMHMQPKKEICALGVCFDEVKEMPQMTLVPAKVASKNVIPALELVEP 301
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 159 -----PLPK----- 184
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 302 IKGEVPAKSDV-----YCEVCEFLVKEVTKLIDNNKTEKEILDAPDKMCKSLPK 351
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 185 LQARPGPHTDLSRQGFPIPLPYCWLCKRALIKRIQAMIPKG-----ALAVAQAQCRVVPL 240
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 352 SLSECEQVVDVTVSSILSLILEVSPVLCVSMHLCS-----CTRLPALTVHVTQPKDG 406
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 241 VAGIGICQCLAEYRVSIVLLDTLGLRMLPOLVCRVLRCMSDDSDAGPRSP---TGSWLPKDS 297
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 407 GFCEVCKKLVGLDRNLKNSKTEKILAALEKGC--SFLPDYKQKCDQFVAEYEPVLIE 464
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 298 E-CHLQNSVT-----TQAGNSSEAIQAMLAQCVGSLD---REKCKQFVEQHTPQLLT 348
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 465 ILVEVMDPSFVCLKIGACPSAKHPLLGTCKIWP 499
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 349 LVPRGWDHITTCQALGCVGTMSPL-----QCIHSP 379
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

RESULT 10
PSPB_CANFA STANDARD; PRT; 363 AA.
AC P17129;
DT 01-AUG-1990 (Rel. 15, Created)
DT 01-AUG-1990 (Rel. 15, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Pulmonary surfactant-associated protein B precursor (SP-B) (6 kDa
DE protein) (Pulmonary surfactant-associated protein B) (Phe)
DE (Pulmonary surfactant-associated protein 18) (SP 18) (Fragment).
GN SFTPB OR SFTPB.
OS Canis familiaris (Dog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.
OX NCBI_TaxID=9615;
RN [1]
RP SEQUENCE FROM N.A., AND SEQUENCE OF 182-211.
RC TISSUE=Lung;
RA MEDLINE=87092398; PubMed=3467361;
RA Hawgood S., Benson B.J., Schilling J., Damm D., Clements J.A.,
RA White R.T.;
RT "Nucleotide and amino acid sequences of pulmonary surfactant protein
RT SP 18 and evidence for cooperation between SP 18 and SP 28-36 in
RT surfactant lipid adsorption."
RL Proc. Natl. Acad. Sci. U.S.A. 84:66-70(1987).
CC -!- FUNCTION: PULMONARY SURFACTANT-ASSOCIATED PROTEINS PROMOTE
CC ALVEOLAR STABILITY BY LOWERING THE SURFACE TENSION AT THE AIR-
CC LIQUID INTERFACE IN THE PERIPHERAL AIR SPACES. SP-B INCREASES
CC THE COLLAPSE PRESSURE OF PALMITIC ACID TO NEARLY 70 MILLINEWTONS
CC PER METER.
CC -!- SUBUNIT: Homodimer; disulfide-linked.
CC -!- SUBCELLULAR LOCATION: Extracellular.
```



Db 185 -----LP-----YCMLCRTLKRIQAMI-----PKGVLAVTVQVCHVPLVVGICOC 228  
Qy 360 VVDYGGSSILSILEEVSPELVCSMLHLS-----GTRLPALTVHVTVQKGGFCEVCKK 414  
Db 229 LGERTVLLDALLGRMLPOLVGLVLRCSHDSAGPALASPSE--WSPQESK--COLQWF 286  
Qy 415 LVGYLDNRNLEKSTKQBIILAALKGC--SFLPDPYQKQCDQFVAEVEPVLBIILVEWDP 472  
Db 287 VT-----TOAGNHSEQATPAIQACLSWLD---RQKCEQFVEQHMPRLQTLASGRDA 338  
Qy 473 SFVCLKIGACPSAHKPLGTKEKINGPSY 501  
Db 339 HTTCQALGACRTTFSPL-----QCIHTPHF 363

RESULT 11  
PSPB RABIT  
ID PSPB RABIT STANDARD; PRT; 370 AA.  
AC P15285;  
DT 01-APR-1990 (Rel. 14, Created)  
DT 01-OCT-1996 (Rel. 34, Last sequence update)  
DT 16-OCT-2001 (Rel. 40, Last annotation update)  
DE Pulmonary surfactant-associated protein B precursor (SP-B) (6 kDa  
DE protein) (Pulmonary surfactant-associated proteolipid SPL(Phe)).  
GN SFTPB OR SFTP3.  
OS Oryctolagus cuniculus (Rabbit).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.  
OX NCBI\_TaxID=9986;  
RN [1]  
RN TISSUE=Lung;  
RN MEDLINE=89228033; PubMed=2469419;  
RA Xu J., Richardson C., Ford C., Spencer T., Li-Juan Y., Mackie G.,  
RA Hammond G., Possmayer F.;  
RT "Isolation and characterization of the cDNA for pulmonary surfactant-  
RT associated protein-B (SP-B) in the rabbit.";  
RL Biochem. Biophys. Res. Commun. 160:325-332(1989).  
RN [2]  
RN SEQUENCE FROM N.A.  
RC STRAIN=New Zealand white;  
RX MEDLINE=95208794; PubMed=7900830;  
RA Margana R.K., Boggaram V.;  
RT "Transcription and mRNA stability regulate developmental and hormonal  
RT expression of rabbit surfactant protein B gene.";  
RL Am. J. Physiol. 268:L481-L490(1995).  
RN [3]  
RN SEQUENCE OF 1-34 FROM N.A.  
RX MEDLINE=96096536; PubMed=8522191;  
RA Luzi P., Anneschi M., Strayer D.S.;  
RT "The upstream region of the SP-B gene: intrinsic promoter activity  
RT and glucocorticoid responsiveness related to a new DNA-binding  
RT protein.";  
RL Gene 165:285-290(1995).  
CC -1- FUNCTION: PULMONARY SURFACTANT-ASSOCIATED PROTEINS PROMOTE  
CC ALVEOLAR STABILITY BY LOWERING THE SURFACE TENSION AT THE AIR-  
CC LIQUID INTERFACE IN THE PERIPHERAL AIR SPACES. SP-B INCREASES  
CC THE COLLAPSE PRESSURE OF PALMITIC ACID TO NEARLY 70 MILLINEWTONS  
CC PER METER.  
CC -1- SUBUNIT: Homodimer; disulfide-linked.  
CC -1- SUBCELLULAR LOCATION: Extracellular.  
CC -1- MISCELLANEOUS: Pulmonary surfactant consists of 90% lipid and 10%  
CC protein. There are 4 surfactant-associated proteins: 2 collagenous,  
CC carbohydrate-binding glycoproteins (SP-A and SP-D) and 2 small  
CC hydrophobic proteins (SP-B and SP-C).  
CC -1- SIMILARITY: Contains 1 saposin A-type domain.  
CC -1- SIMILARITY: Contains 3 saposin B-type domains.  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial

entities requires a license agreement (see <http://www.isb-sib.ch/announce/>)  
or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
-----  
CC EMBL; M24901; AAA31466.1; -;  
CC EMBL; U17106; AAG67934.1; -;  
DR EMBL; S80649; AAD14335.1; -;  
DR PIR; A32421; LNRBB.  
DR PIR; I46531; I46531.  
DR HSSP; P07988; IDFW.  
DR InterPro; IPR003119; Sapa.  
DR InterPro; IPR007856; Sapa\_1.  
DR InterPro; IPR008138; Sapa\_2.  
DR InterPro; IPR008140; Sapa\_sub.  
DR InterPro; IPR008139; SaposinB.  
DR InterPro; IPR008137; Surfactant\_B.  
DR Pfam; PF02199; SAPA; 1.  
DR Pfam; PF05184; Sapa\_1; 1.  
DR Pfam; PF03489; Sapa\_2; 3.  
DR ProDom; PD001732; Sapa\_sub; 1.  
DR ProDom; PD008002; Surfactant\_B; 1.  
DR SMART; SM00162; SAPA; 1.  
DR SMART; SM00118; SAPP; 3.  
KW Surface film; Gaseous exchange; Glycoprotein; Repeat.  
FT PROPEP 1 184  
FT CHAIN 185 263 PULMONARY SURFACTANT-ASSOCIATED PROTEIN  
FT PROPEP 264 370 B.  
FT DOMAIN 29 62 SAPOSIN-LIKE TYPE A.  
FT DISULFID 192 261 BY SIMILARITY.  
FT DISULFID 195 255 BY SIMILARITY.  
FT DISULFID 219 230 BY SIMILARITY.  
FT DISULFID 232 232 INTERCHAIN (BY SIMILARITY).  
FT CARBOHYD 300 300 N-LINKED (GLCNAC...) (POTENTIAL).  
FT CONFLICT 15 15 MISSING (IN REF. 2).  
FT CONFLICT 184 184 R -> P (IN REF. 1).  
FT CONFLICT 232 232 C -> R (IN REF. 2).  
FT CONFLICT 329 355 ELHTPOLLSLSRGWDARAICQALGAC -> AHAHAAAEPA  
FT SEQUENCE 370 AA; 40610 MW; 423047A69B12DCB5 CRC64;  
Query Match 10.6%; Score 296.5; DB 1; Length 370;  
Best Local Similarity 21.7%; Pred. No. 1.7e-12;  
Matches 111; Conservative 65; Mismatches 170; Indels 165; Gaps 18;  
Qy 4 LFLASLLGAALAGPVLGKECTGSAVVCQNVTSADCGAVKCHLQTVNKNPTVKSLPC 63  
Db 12 LLLLTLCGPGTAVWATSPACAGPEFWCQSEALQCKALGCHQEVWGHVGADDL-C 70  
Qy 64 DICKDVTAAAGDMLKDNATEEILVYLEKTCMDLWPKENMSASCKEIVDSYLPVLDIIG 123  
Db 71 QECQDIVNLTKEAIFQDTIRKLEHEDVLPLKLLVPOCHHVLVDVFPITITVFQS 130  
Qy 124 EMSRGEVCSALNLCESLQKHLAELNHQKLESNKIPELDMTEVAVFPMANIPLLLYPD 183  
Db 131 QINAKA-ICQHLGLCQ-----PGSPEPLD-----PLPDKLVLPPLL--G 167  
Qy 184 GPRSPKOPKONGDVQCQCIQVMTDIQTAVRTNSTFVQALVEHVKEECRDPGMAICKN 243  
Db 168 ALPAKP-----GPHQTQJUS-- 181  
Qy 244 YISQYSEIAIQMMHMPKPEICALVGFCDVEKEMPQTLPAKVASKNVIPALEVEPIK 303  
Db 182 -----AQRFPPL-----PL- 191  
Qy 304 KHEVPAKSDVYCEVCEPLVKEVTKLIDNNKTEKELD-AFDKMCCKLPKLSLSECEQVVD 362  
Db 192 -----CWLRTLLKRIQAM-----PGVLAMAVAQVCHVPLVVGICQCLAE 235  
Qy 363 TYGSSILSILEEVSPELVCSMLHLS-----GTRLPALTVHVTVQ--PKDGGFCEVCKK 415  
Db 236 RYTVILLVLLGHVLPQVGLVLRCSVDISIQVFPPTLEALPGEWLPDPPE-CRLCMSV 294  
Qy 416 VGVLDNRNLEKSTKQBIILAALKGC-----LEKGCSPDPYQKQCDQFVAEVEPVLBIILVEYM 470



```

Db 295 TQAR-RNISEQTRQAVYHACLSQLDK-----QCEQFVELHTPQLLSLRGW 343
Oy 471 DPSVFLKIGACPSAHKPLLTGKCMGPSY 501
Db 344 DARAICQALGACVATLSPL-----QCISPHF 370

RESULT 12
SAP_CAVPO
ID SAP_CAVPO STANDARD; PRT; 81 AA.
AC P20097;
DT 01-FEB-1991 (Rel. 17, Created)
DT 01-FEB-1991 (Rel. 17, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Saposin C (CO-beta-glucosidase) (Glucosylceramide activator)
DE (Sphingolipid activator protein 2) (SAP-2).
OS Cavia porcellus (Guinea pig).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Hystricognathi; Caviidae; Cavia.
OX NCBI_TaxID=10141;
RN [1]
RP SEQUENCE.
RC TISSUE=Liver;
RX MEDLINE=89066787; PubMed=3198642;
RA Sano A., Radin N.S., Johnson L.L., Tarr G.E.;
RT "The activator protein for glucosylceramide beta-glucosidase from
RT guinea pig liver. Improved isolation method and complete amino acid
RT sequence".
RL J. Biol. Chem. 263:19597-19601(1988).
CC -!- FUNCTION: Saposin A and saposin C stimulate the hydrolysis of
CC glucosylceramide by beta-glucosylceramidase (EC 3.2.1.45) and
CC galactosylceramide by beta-galactosylceramidase (EC 3.2.1.46).
CC Saposin C apparently acts by combining with the enzyme and acidic
CC lipid to form an activated complex, rather than by solubilizing
CC the substrate.
CC -!- SIMILARITY: Contains 1 saposin B-type domain.
DR PIR; A32026; A32026.
DR HSSP; P42210; IQDM.
DR InterPro; IPR007856; SapB_1.
DR InterPro; IPR008138; SapB_2.
DR InterPro; IPR008140; SapB_sub.
DR InterPro; IPR008373; Saposin.
DR InterPro; IPR008139; SaposinB.
DR Pfam; PF05184; SapB_1; 1.
DR Pfam; PF03489; SapB_2; 1.
DR PRINTS; PR01797; SAPOSIN.
DR ProDom; PD001732; SapB_sub; 1.
DR SMART; SM00118; SAPB; 1.
KW Glycoprotein; Sphingolipid metabolism.
FT DOMAIN 1 81 SAPOSIN-LIKE TYPE B.
FT CARBOHYD 22 22 N-LINKED (GLCNAC. . .).
FT DISULFID 5 78 BY SIMILARITY.
FT DISULFID 8 72 BY SIMILARITY.
FT DISULFID 36 47 BY SIMILARITY.
SQ SEQUENCE 81 AA; 8852 MW; E564CE1FOA292596 CRC64;

Query Match 9.4%; Score 261; DB 1; Length 81;
Best Local Similarity 61.8%; Pred. No. 5.5e-11;
Matches 47; Conservative 18; Mismatches 11; Indels 0; Gaps 0;

Oy 313 VYCVCEFLVKEVTKLIDNNKTEKILDAFKMGSKLPKLSSECEQVVDYTGSSILSIL 372
Db 3 VTKACEVYVVKVWELIDNNTEKEIKIHALDSVALLPESVSEVCEQVVDYTGDSIVALL 62

Oy 373 LEEVSPELVCSMLHLC 388
Db 63 LQEMSPELVCSSELGLC 78

RESULT 13
ASPR_CUCPE
ID ASPR_CUCPE STANDARD; PRT; 513 AA.

```

```

AC 004057;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Aspartic proteinase precursor (EC 3.4.23.-).
OS Cucurbita pepo (Vegetable marrow) (Summer squash).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OX eurosids I; Cucurbitales; Cucurbitaceae; Cucurbita.
OX NCBI_TaxID=3663;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=97354181; PubMed=9210475;
RA Hiraiwa N., Kondo M., Nishimura M., Hara-Nishimura I.;
RT "An aspartic endopeptidase is involved in the breakdown of
RT propeptides of storage proteins in protein-storage vacuoles of
RT plants.".
RL Eur. J. Biochem. 246:133-141(1997).
CC -!- FUNCTION: Involved in the breakdown of propeptides of storage
CC proteins in protein-storage vacuoles.
CC -!- SUBUNIT: Heterodimer of a 32 kDa subunit and a 16 kDa subunit.
CC -!- SUBCELLULAR LOCATION: Vacuolar.
CC -!- SIMILARITY: Belongs to peptidase family A1.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC -----
DR EMBL; AB002695; BAA19607.1; -
DR PIR; T09739; T09739.
DR HSSP; P42210; IQDM.
DR MEROPS; A01.020; -.
DR InterPro; IPR001969; Asprotease_AS.
DR InterPro; IPR009007; Pept_A_acid.
DR InterPro; IPR001461; Peptidase_A1.
DR InterPro; IPR007856; SapB_1.
DR InterPro; IPR008138; SapB_2.
DR InterPro; IPR008140; SapB_sub.
DR InterPro; IPR008373; Saposin.
DR InterPro; IPR008139; SaposinB.
DR Pfam; PF00026; asp; 1.
DR Pfam; PF05184; SapB_1; 1.
DR Pfam; PF03489; SapB_2; 1.
DR PRINTS; PR00792; PEPSIN.
DR PRINTS; PR01797; SAPOSIN.
DR ProDom; PD001732; SapB_sub; 1.
DR SMART; SM00118; SAPB; 2.
DR PROSITE; PS00141; ASP_PROTEASE; 2.
KW Hydrolase; Aspartyl protease; Zymogen; Glycoprotein; Signal.
FT SIGNAL 1 25 POTENTIAL.
FT PROPEP 26 ? POTENTIAL.
FT CHAIN ? 513 ASPARTIC PROTEINASE.
FT DOMAIN 322 420 SPECIFIC TO PLANT ASPARTIC PROTEINASES.
FT ACT_SITE 108 108 BY SIMILARITY.
FT ACT_SITE 294 294 BY SIMILARITY.
FT DISULFID 120 126 BY SIMILARITY.
FT DISULFID 285 289 BY SIMILARITY.
FT DISULFID 432 469 BY SIMILARITY.
FT CARBOHYD 143 143 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 404 404 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 513 AA; 55855 MW; 083FB7064CE02DC2 CRC64;

Query Match 6.2%; Score 173.5; DB 1; Length 513;
Best Local Similarity 29.0%; Pred. No. 0.00031;
Matches 47; Conservative 31; Mismatches 59; Indels 25; Gaps 7;

Oy 353 LSECEQVVDYTGSSILSILLESVSPELVCSMLHLC--GTRLPALTVH-VTPKDG--- 406
Db 320 VSQCKAVVAQYGTQIMDLLSEADPKKISQINLCTFDGTRGVSMGTESVVDENAGKSS 379

```

```
Oy 407 -----GFEYCKLVGLDRNLEKSTKQELLALEKSGSFLPDPYQK---QCQDFVAEY 458
Db 380 DSLHDGNCSCVEMTVVMQNLQKQTKERININELCDRPMSPMGQSAVDCGO-LSSM 438
Oy 459 EPLVIELVEMVD--PSFVCLKIGACPSAKHPLLGTEKIWG 498
Db 439 PTVSFTIGGKIFDLAPEEYILKVGEGVA-----QCISG 472

RESULT 14
ID ASRP_HORVU STANDARD; PRT; 508 AA.
AC P42210;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE Phytapsin precursor (RC 3.4.23.40) (Aspartic proteinase).
OS Hordeum vulgare (Barley).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Poideae;
OC Triticeae; Hordeum.
OX NCBI_TaxID=4513;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Kustaa; TISSUE=Embryo;
RX MEDLINE=92111473; PubMed=1722454;
RA Runeberg-Roos P., Toermaekangas K., Oestman A.;
RT "Primary structure of a barley-grain aspartic proteinase. A plant
aspartic proteinase resembling mammalian cathepsin D.";
RL Eur. J. Biochem. 202:1021-1027(1991).
RN [2]
RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
RC STRAIN=cv. Kustaa; TISSUE=Embryo;
RX MEDLINE=92254717; PubMed=1812727;
RA Toermaekangas K., Runeberg-Roos P., Oestman A., Tilgmann C.,
RA Sarkkinen P., Kervinen J., Mikola L., Kalkkinen N.;
RT "Aspartic proteinase from barley seeds is related to animal cathepsin
D.";
RL Adv. Exp. Med. Biol. 306:355-359(1991).
RN [3]
RP X-RAY CRYSTALLOGRAPHY (2.3 ANGSTROMS).
RX MEDLINE=99335466; PubMed=10406799;
RA Kervinen J., Tobin G.J., Costa J., Waugh D.S., Wlodawer A., Zdanov A.;
RT "Crystal structure of plant aspartic proteinase phytypsin:
inactivation and vacuolar targeting.";
RL EMBO J. 18:3947-3955(1999).
CC -! CATALYTIC ACTIVITY: Prefers hydrophobic residues Phe, Val, Ile,
Leu, and Ala at P1 and P1', but also cleaves -Phe-|-Asp- and
-Asp-|-Asp- bonds in 2S albumin from plants seeds.
CC -! SUBUNIT: Heterodimer of two subunits (29 kDa and 11 kDa) processed
from the precursor molecule. A large enzyme (32 kDa and 16 kDa) is
an intermediate precursor form.
CC -! SUBCELLULAR LOCATION: Vacuolar.
CC -! TISSUE SPECIFICITY: Embryo and leaf.
CC -! SIMILARITY: Belongs to peptidase family A1.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (see http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
CC -----
CC EMBL; X56136; CAA39602.1; -
CC PIR; S19697; S19697.
CC PDB; 1QDM; 16-JUL-99.
CC MEROPS; A01.020; -.
CC InterPro; IPR001969; Asprotease AS.
CC InterPro; IPR009007; Pept A acid.
CC InterPro; IPR001461; Peptidase_A1.
CC InterPro; IPR007856; SapB_1.
```

```
DR InterPro; IPR008138; SapB_2.
DR InterPro; IPR008140; SapB_sub.
DR InterPro; IPR008173; Saposin.
DR InterPro; IPR008139; SaposinB.
DR Pfam; PF00026; asp; 1.
DR Pfam; PF05184; SapB_1; 1.
DR Pfam; PF03489; SapB_2; 1.
DR PRINTS; PR00792; PEPSIN.
DR PRINTS; PR01797; SAPOSIN.
DR ProDom; PD001732; SapB_sub; 1.
DR SMART; SM00118; SAPB; 2.
DR PROSITE; PS00141; ASP_PROTEASE; 2.
KW Hydrolase; Aspartyl protease; Zymogen; Glycoprotein; Signal;
KW 3D-structure.
FT SIGNAL 1 22 POTENTIAL.
FT PROPEP 23 66
FT CHAIN 67 377 PHYTEPSIN 32 kDa SUBUNIT.
FT CHAIN 67 377 PHYTEPSIN 29 kDa SUBUNIT.
FT CHAIN 378 508 PHYTEPSIN 16 kDa SUBUNIT.
FT CHAIN 422 508 PHYTEPSIN 11 kDa SUBUNIT.
FT DOMAIN 317 420 SPECIFIC TO PLANT ASPARTIC PROTEINASES.
FT SITE 377 378 CLEAVAGE.
FT SITE 421 422 CLEAVAGE.
FT ACT_SITE 102 102
FT ACT_SITE 289 289
FT DISULFID 115 121
FT DISULFID 280 284
FT DISULFID 427 464
FT CARBOHYD 399 399
FT STRAND 32 37
FT HELIX 42 49
FT TURN 50 50
FT HELIX 74 76
FT STRAND 79 81
FT TURN 84 90
FT TURN 91 94
FT STRAND 95 102
FT TURN 103 104
FT STRAND 109 112
FT TURN 113 114
FT STRAND 117 123
FT STRAND 129 127
FT HELIX 133 134
FT TURN 133 134
FT STRAND 136 136
FT STRAND 141 146
FT TURN 147 148
FT STRAND 149 162
FT TURN 163 164
FT STRAND 165 177
FT STRAND 180 180
FT HELIX 182 185
FT STRAND 190 193
FT HELIX 197 199
FT HELIX 201 203
FT HELIX 207 211
FT TURN 212 215
FT STRAND 221 225
FT STRAND 236 240
FT TURN 241 241
FT TURN 245 246
FT STRAND 248 260
FT TURN 261 262
FT STRAND 263 267
FT STRAND 270 272
FT TURN 273 274
FT STRAND 275 276
FT TURN 279 282
FT STRAND 284 288
FT STRAND 295 297
FT HELIX 299 309
FT TURN 310 310
FT STRAND 314 314
```

N-LINKED (GLCNAC. . .) (POTENTIAL).

```

FT HELIX 317 325
FT TURN 326 326
FT HELIX 327 335
FT TURN 336 337
FT HELIX 340 346
FT TURN 347 348
FT HELIX 379 397
FT TURN 398 399
FT HELIX 402 412
FT TURN 413 414
FT STRAND 416 416
FT STRAND 424 425
FT HELIX 427 430
FT TURN 431 432
FT STRAND 436 440
FT TURN 441 442
FT STRAND 443 447
FT HELIX 449 452
FT STRAND 453 455
FT HELIX 459 461
FT STRAND 464 466
FT STRAND 468 470
FT TURN 475 476
FT STRAND 480 483
FT HELIX 485 488
FT TURN 489 490
FT STRAND 491 496
FT TURN 497 500
FT STRAND 501 508
SQ SEQUENCE 508 AA; 54226 MW; 87F2C9F93369B962 CRC64;

Query Match 5.7%; Score 158.5; DB 1; Length 508;
Best Local Similarity 24.3%; Pred. No. 0.003;
Matches 49; Conservative 40; Mismatches 78; Indels 35; Gaps 9;

Qy 193 DNGDV-----CQCIQWTDIQAVRTNSTFVQALVHVKECDRLPGMADICKN 243
Db 266 DMGDLVGGKSTGFCAGCAIADSGTSLLAGPT---AIITEINEKICAGGV-VSQECKT 321
Qy 244 YISQYSEIAIQMMW-HMQPKIEICALGVCFDEKEMPMQTLVPAKVASKNVPALVELPEI 302
Db 322 IVSYGGQILLDAETQPKKICQVGLC-----TFDGTGVSAGIRSVVD-DEPV 371
Qy 303 KHQEVPAKSDYVCVCFVKEVTKLIDNNKTEKILDAFDKMGSKLPKSLSE---RC-- 357
Db 372 KSNGL--RADPMCSACEMAVVMQNLQAQNKTDLLDYNQLCNRLPSPMGESAVDCGS 429
Qy 358 -----QEVVDYVGSSILSILEE 375
Db 430 LGSMPDIEFTIGGKFKPKPEE 451

```

## RESULT 15

```

CYP1_CYNCA STANDARD; PRT; 473 AA.
AC P40782;
DT 01-FEB-1995 (Rel. 31, Created)
DT 01-FEB-1995 (Rel. 32, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Cyprosin precursor (EC 3.4.23.-) (Fragment).
GN CYP101.
OS Cynara cardunculus (Cardoon).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; asterids;
OC campanulids; Asterales; Asteraceae; Carduoideae; Cardueae; Cynara.
OK NCBI_TaxID=4265;
RN [1]
RP SEQUENCE FROM N.A., AND SEQUENCE OF 178-186.
RC TISSUE=Flower buds;
RX MEDLINE=94250836; PubMed=8193298;
RA Cordeiro M.C.; Xue Z.-T.; Pietrzak M.; Pais M.S.; Brodelius P.E.;
RT "Isolation and characterization of a cDNA from flowers of Cynara
cardunculus encoding cyprosin (an aspartic proteinase) and its use to

```

```

RT study the organ-specific expression of cyprosin.";
RL Plant Mol. Biol. 24:733-741(1994).
CC -I- TISSUE SPECIFICITY: Mostly present in the violet parts of styles
CC and corollas of mature flowers.
CC -I- DEVELOPMENTAL STAGE: Expressed in early stages of floral
CC development and switched off at maturation of the flower.
CC -I- SIMILARITY: Belongs to peptidase family A1.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC -----
DR EMBL; X69193; CAA48939.1; ALT_SEQ.
DR HSSP; P42210; IQDM.
DR MEROPS; A01.020; -.
DR InterPro; IPR001969; Aspprotease AS.
DR InterPro; IPR009007; Pept A acid.
DR InterPro; IPR001461; Peptidase_A1.
DR InterPro; IPR007856; SapB_1.
DR InterPro; IPR008138; SapB_2.
DR InterPro; IPR008140; SapB sub.
DR InterPro; IPR008139; SaposinB.
DR Pfam; PF00026; asp; 1.
DR Pfam; PF05184; SapB_1; 1.
DR Pfam; PF03489; SapB_2; 1.
DR PRINTS; PR00792; PEPFIS.
DR ProDom; PD001732; SapB sub; 1.
DR SMART; SM00118; SAPB; 2.
DR PROSITE; PS00141; ASP_PROTEASE; 2.
KW Hydrolase; Aspartyl protease; Zymogen; Glycoprotein.
FT NON_TER 1
FT PROPEP <1 33
FT CHAIN 34 473
FT ACT_SITE 69 69
FT ACT_SITE 256 256
FT DISULFID 82 88
FT DISULFID 247 251
FT DISULFID 392 429
FT CARBOHYD 364 364
FT CARBOHYD 411 411
SQ SEQUENCE 473 AA; 51564 MW; 65F3232EBD06CB56 CRC64;

Query Match 5.5%; Score 154.5; DB 1; Length 473;
Best Local Similarity 30.1%; Pred. No. 0.005;
Matches 31; Conservative 24; Mismatches 39; Indels 9; Gaps 2;

Qy 353 LSEECOEVDVDTYGSILSILLESVSPELVCSMLHLCS--GTRLPALTVHVTQPKDGG--- 407
Db 282 MSQCKSLVDVQYQKSMIEMLLSEQPEKICQMKLCSFDGSHDTSMLIESVVDKSKGSS 341
Qy 408 ----FCEVCKKLYGILDRNLERNKSTKQEIILAALEKGCSPFLDP 446
Db 342 GLPMRCVPCARWVVMQNIQRQNETENIINYVDKLCERLPS 384

Search completed: May 5, 2004, 13:31:42
Job time : 14.9383 secs

```

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 5, 2004, 13:26:23 ; Search time 20.9003 Seconds  
(without alignments)  
2411.659 Million cell updates/sec

Title: US-09-743-684A-1  
Perfect score: 2789  
Sequence: 1 MYALFLLASLLGALAGPVL.....NTETAQCNVAECHKRHVMN 524  
Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 283366 seqs, 96191526 residues

Total number of hits satisfying chosen parameters: 283366

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : PIR 78:\*  
1: pir1:\*  
2: pir2:\*  
3: pir3:\*  
4: pir4:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2777.5	99.6	527	1 SAHUP	saposin precursor
2	1996	71.6	554	1 A28716	saposin precursor
3	1937.5	69.5	557	1 JH0604	saposin precursor
4	511	18.3	965	2 T00207	p109 protein - sil
5	389	13.9	79	2 A49475	cerebroside sulfat
6	379	13.6	376	2 S02766	pulmonary surfacta
7	348.5	12.5	381	1 LNHUB	pulmonary surfacta
8	343	12.3	80	2 S21770	saposin-C - bovine
9	310	11.1	363	2 A29072	pulmonary surfacta
10	301.5	10.8	369	2 I46531	surfactant protein
11	261	9.4	81	2 A32036	glucosylceramide b
12	256.5	9.2	370	1 LNRBB	pulmonary surfacta
13	213	7.6	213	2 T46069	hypothetical prote
14	203.5	7.3	217	2 T48201	hypothetical prote
15	176.5	6.3	402	2 T15677	hypothetical prote
16	173.5	6.2	513	2 T09739	aspartic endopepti
17	171.5	6.1	506	2 F86253	hypothetical prote
18	165.5	5.9	505	2 T07916	probable aspartic
19	161.5	5.8	513	2 T11686	aspartic proteinas
20	161	5.8	428	2 S47096	cynarase (EC 3.4.2
21	161	5.8	474	2 T12049	cyprosin (EC 3.4.2
22	158.5	5.7	508	2 S19697	aspartic proteinas
23	156	5.6	314	2 T15674	hypothetical prote
24	155.5	5.6	292	2 T14446	aspartic proteinas
25	155.5	5.6	322	2 S41400	aspartic proteinas
26	154	5.5	496	2 JS0732	aspartic proteinas
27	153.5	5.5	509	2 JC7272	aspartic proteinas
28	145.5	5.2	205	2 B89567	protein T08A9.7 (i
29	143.5	5.1	1175	2 S52417	E-selectin ligand-

ALIGNMENTS

RESULT 1

SAHUP

saposin precursor [validated] - human

N/Alternate names: cerebroside sulfate activator protein; co-beta-glucosidase; component ein (SAP); sphingolipid activator protein A2; sulfate sulfate activator protein  
N/Contains: prosaposin; saposin A; saposin B; saposin C; saposin D  
C/Species: Homo sapiens (man)  
C/Date: 30-Jun-1992 #sequence revision 17-Nov-1995 #text change 08-Dec-2000  
C/Accession: JX0061; A57368; A42003; B42003; C42003; D42003; A30367; S34740; S36140; S36  
0226; I37265; I37264  
R/Nakano, T.; Sandhoff, K.; Stuenkel, J.; Christomanou, H.; Suzuki, K.  
J. Biochem. 105, 152-154, 1989  
A/Title: Structure of full-length cDNA coding for sulfate activator, a Co-beta-glucosid  
A/Reference number: JX0061; MUID:89255151; PMID:2498298  
A/Accession: JX0061  
A/Molecule type: mRNA  
A/Residues: 1-527 <NAK>  
A/Cross-references: GB:D00422; NID:g220063; PIDN:BAA00321.1; PID:g220064  
A/Note: alternative splice form 1  
A/Accession: A57368  
A/Molecule type: mRNA  
A/Residues: 1-259,263-527 <NA2>  
A/Cross-references: GB:J03015; GB:J03086; NID:G337755; PIDN:AAB59494.1; PID:G337756  
A/Note: alternative splice form 2  
R/Roman, E.G.; Scheinker, V.; Grabowski, G.A.  
Genomics 13, 312-318, 1992  
A/Title: Structure and evolution of the human prosaposin chromosomal gene.  
A/Reference number: A42003; MUID:92307663; PMID:1612590  
A/Accession: A42003  
A/Molecule type: DNA  
A/Residues: 50-140 <ROR>  
A/Cross-references: GB:M86181  
A/Note: sequence extracted from NCBI backbone (NCBIN:107235, NCBIP:107236)  
A/Accession: B42003  
A/Molecule type: DNA  
A/Residues: 185-259,263-276 <RO2>  
A/Note: sequence extracted from NCBI backbone (NCBIN:107235, NCBIP:107237)  
A/Accession: C42003  
A/Molecule type: DNA  
A/Residues: 305-393 <RO3>  
A/Note: sequence extracted from NCBI backbone (NCBIN:107235, NCBIP:107238); sequence inc  
A/Accession: D42003  
A/Molecule type: DNA  
A/Residues: 399-487 <RO4>  
A/Note: sequence extracted from NCBI backbone (NCBIN:107235, NCBIP:107239); sequence inc  
R/Roman, E.G.; Grabowski, G.A.  
Genomics 5, 486-492, 1989  
A/Title: Molecular cloning of a human co-beta-glucosidase cDNA: evidence that four spin  
A/Reference number: A30367; MUID:90129043; PMID:2515150  
A/Accession: A30367  
A/Molecule type: mRNA  
A/Residues: 1-259,263-527 <RO5>

A;Cross-references: GB:J03077; NID:G183230; PIDN:AAA52560.1; PID:G183231  
A;Note: alternative splice form 2  
R;Hiraiwa, M.; O'Brien, J.S.; Kishimoto, Y.; Galdick, M.; Fluharty, A.L.; Ginns, E.I.; Arch. Biochem. Biophys. 304, 110-116, 1993  
A;Title: Isolation, characterization, and proteolysis of human prosaposin, the precursor of saposin A  
A;Reference number: S34740; MUID:93311991; PMID:8323276  
A;Accession: S34740  
A;Molecule type: protein  
A;Residues: 17-24;165-172;180-189;301-305 <HIR>  
R;Tyynelae, J.; Palmer, D.N.; Baumann, M.; Haltia, M. FEBS Lett. 330, 8-12, 1993  
A;Title: Storage of saposins A and D in infantile neuronal ceroid-lipofuscinosis.  
A;Reference number: S36140; MUID:93380576; PMID:8370464  
A;Accession: S36140  
A;Molecule type: protein  
A;Residues: 'XX', 62, 'X', 64-65, 'X', 67-79, 'X', 81-84 <TY>  
A;Note: saposin A  
A;Accession: S36141  
A;Molecule type: protein  
A;Residues: 'XX', 413-414, 'X', 416-428, 'X', 430-434 <TV2>  
A;Note: saposin D  
R;Holtzman, H.; Sandhoff, K.; Kwon, H.Y.; Harzer, K.; Nakano, T.; Suzuki, K. J. Biol. Chem. 266, 7556-7560, 1991  
A;Title: Sulfatide activator protein. Alternative splicing that generates three mRNAs and a protein  
A;Reference number: S36988; MUID:91210267; PMID:2019586  
A;Accession: S36988  
A;Status: nucleic acid sequence not shown; translation not shown  
A;Molecule type: mRNA  
A;Residues: 1-240, 'S', 242-527 <HOL>  
A;Cross-references: EMBL:M60255; NID:G337759; PIDN:AAA36594.1; PID:G337760  
A;Note: the nucleotide sequence was submitted to the EMBL Data Library, June 1991  
A;Note: cerebroside sulfate activator protein mutant MU-9; corresponds to alternative splicing  
A;Accession: S36989  
A;Status: nucleic acid sequence not shown; translation not shown  
A;Molecule type: mRNA  
A;Residues: 1-240, 'S', 242-529, 263-527 <HOL>  
A;Cross-references: EMBL:M60257; NID:G337764; PIDN:AAA36595.1; PID:G337765  
A;Note: the nucleotide sequence was submitted to the EMBL Data Library, June 1991  
A;Note: cerebroside sulfate activator protein mutant MU-0; corresponds to alternative splicing  
A;Accession: S36990  
A;Status: nucleic acid sequence not shown; translation not shown  
A;Molecule type: mRNA  
A;Residues: 1-240, 'S', 242-259, 261-527 <HOL>  
A;Cross-references: EMBL:M60258; NID:G337766; PIDN:AAA36596.1; PID:G337767  
A;Note: the nucleotide sequence was submitted to the EMBL Data Library, June 1991  
A;Note: cerebroside sulfate activator protein mutant MU-6; corresponds to alternative splicing  
R;Kondoh, K.; Hinenio, T.; Sano, A.; Kakimoto, Y. Biochem. Biophys. Res. Commun. 181, 286-292, 1991  
A;Title: Isolation and characterization of prosaposin from human milk.  
A;Reference number: PS0330; MUID:92068206; PMID:1958198  
A;Accession: PS0330  
A;Molecule type: protein  
A;Residues: 17-24, 'X', 26 <KON>  
A;Experimental source: milk  
R;Kretz, K.A.; Carson, G.S.; Morimoto, S.; Kishimoto, Y.; Fluharty, A.L.; O'Brien, J.S. Proc. Natl. Acad. Sci. U.S.A. 87, 2541-2544, 1990  
A;Title: Characterization of a mutation in a family with saposin B deficiency: a glycosylation defect  
A;Reference number: A35985; MUID:90207231; PMID:2320574  
A;Accession: A35985  
A;Molecule type: mRNA  
A;Residues: 213-221 <KRE>  
A;Cross-references: GB:M32221  
A;Accession: B35985  
A;Status: nucleic acid sequence not shown; translation not shown  
A;Molecule type: mRNA  
A;Residues: 1-259, 263-527 <KR2>  
A;Cross-references: GB:M32221; NID:G337761; PIDN:AAA60303.1; PID:G337762  
A;Experimental source: lymphoblast  
A;Accession: C35985  
A;Molecule type: mRNA  
A;Residues: 213-216, 'I', 218-221 <KR3>  
A;Note: sequence from patients with activator-deficient metachromatic leukodystrophy: the R;Fuerst, W.; Schubert, J.; Machleidt, W.; Meyer, H.E.; Sandhoff, K.

Eur. J. Biochem. 192, 709-714, 1990  
A;Title: The complete amino-acid sequences of human ganglioside GM2 activator protein and its precursor  
A;Reference number: S13195; MUID:91006185; PMID:2209618  
A;Accession: S13196  
A;Molecule type: protein  
A;Residues: 195-259, 263-277 <FUE>  
R;Morimoto, S.; Martin, B.M.; Yamamoto, Y.; Kretz, K.A.; O'Brien, J.S.; Kishimoto, Y. Proc. Natl. Acad. Sci. U.S.A. 86, 3389-3393, 1989  
A;Title: Saposin A: second cerebrosidase activator protein.  
A;Reference number: A32784; MUID:89240739; PMID:2717620  
A;Accession: A32784  
A;Molecule type: protein  
A;Residues: 60-84;86-107;109-119;125-134 <MOR>  
R;O'Brien, J.S.; Kretz, K.A.; Dewji, N.; Wenger, D.A.; Esch, F.; Fluharty, A.L. Science 241, 1098-1101, 1988  
A;Title: Coding of two sphingolipid activator proteins (SAP-1 and SAP-2) by same genetic information  
A;Reference number: A41240; MUID:88321660; PMID:2842863  
A;Accession: A41240  
A;Molecule type: mRNA  
A;Residues: 'GSSR', 18-259, 263-299, 'D', 301-302, 'D', 304-527 <OAB>  
A;Cross-references: GB:J03086  
R;Dewji, N.N.; Wenger, D.A.; O'Brien, J.S. Proc. Natl. Acad. Sci. U.S.A. 84, 8652-8656, 1987  
A;Title: Nucleotide sequence of cloned cDNA for human sphingolipid activator protein 1 p1  
A;Reference number: S02289; MUID:88068647; PMID:2825202  
A;Accession: S02289  
A;Status: significant sequence differences  
A;Molecule type: mRNA  
A;Cross-references: EMBL:J03015  
A;Note: this sequence corrected by A41240  
R;Kleinschmidt, T.; Christomanou, H.; Braunitzer, G. Biol. Chem. Hoppe-Seyler 369, 1361-1365, 1988  
A;Title: Complete amino-acid sequence of the naturally occurring A(2) activator protein 1 p1  
A;Reference number: S02028; MUID:89207118; PMID:3242555  
A;Accession: S02028  
A;Molecule type: protein  
A;Residues: 195-259, 263-276 <KLE>  
R;Fuerst, W.; Machleidt, W.; Sandhoff, K. Biol. Chem. Hoppe-Seyler 369, 317-328, 1988  
A;Title: The precursor of sulfatide activator protein is processed to three different products  
A;Reference number: S00813; MUID:89000190; PMID:3048308  
A;Accession: S00813  
A;Molecule type: protein  
A;Residues: 410-487 <PU2>  
R;Kleinschmidt, T.; Christomanou, H.; Braunitzer, G. Biol. Chem. Hoppe-Seyler 368, 1571-1578, 1987  
A;Title: Complete amino-acid sequence and carbohydrate content of the naturally occurring A(2) activator protein 1 p1  
A;Reference number: S00226; MUID:88163077; PMID:3442600  
A;Accession: S00226  
A;Molecule type: protein  
A;Residues: 314-393 <KL2>  
R;Vacarro, A.M.; Salvio, R.; Barca, A.; Tatti, M.; Ciaffoni, F.; Maras, B.; Siciliano, J. Biol. Chem. 270, 9953-9960, 1995  
A;Title: Structural analysis of saposin C and B. Complete localization of disulfide bridge  
A;Reference number: A57297; MUID:95247790; PMID:7730378  
R;Kleinschmidt, H.; Sandhoff, K.; Fuerst, W.; Kwon, H.Y.; Schnabel, D.; Suzuki, K. FEBS Lett. 280, 267-270, 1991  
A;Title: The organization of the gene for the human cerebrosidase sulfate activator protein  
A;Reference number: I37264; MUID:91192146; PMID:2013321  
A;Accession: I37265  
A;Status: translated from GB/EMBL/DBJ  
A;Molecule type: DNA  
A;Residues: 59-125 <RES>  
A;Cross-references: EMBL:X57107; NID:G30234; PIDN:CAA40391.1; PID:G30235  
A;Accession: I37264  
A;Status: translated from GB/EMBL/DBJ  
A;Molecule type: DNA  
A;Residues: 307-516 <RE2>  
A;Cross-references: EMBL:X57108; NID:G30232; PIDN:CAA40392.1; PID:G1565257  
A;Note: sequence revised relative to PID:G30233 (corrected coding region)  
C;Genetics:

A;Gene: GDB:PSAP; GLBA  
A;Cross-references: GDB:120366; OMIM:176801  
A;Map position: 10q22.1-10q22.1  
A;Introns: 83/3; 338/3; 401/1; 453/3; 480/3  
A;Note: defects in this gene may cause variant Gaucher disease, variant Tay-Sachs disease  
A;Note: list of introns is incomplete

Query Match 99.6%; Score 2777.5; DB 1; Length 527;  
Best Local Similarity 99.4%; Pred. No. 1.1e-175;  
Matches 524; Conservative 0; Mismatches 0; Indels 3; Gaps 1;

Qy 1 MYALFLASLLGAALGAPVLGLKECTRGSAVWCNQVKTASDCGAVKHCLQTVWNKPTVKS 60  
Db 1 MYALFLASLLGAALGAPVLGLKECTRGSAVWCNQVKTASDCGAVKHCLQTVWNKPTVKS 60  
Qy 61 LPCDICVDVTAAGDMLKDNATEEILVLEKTCDWLPKPNMSASCKEIVDSYLPVILDI 120  
Db 61 LPCDICVDVTAAGDMLKDNATEEILVLEKTCDWLPKPNMSASCKEIVDSYLPVILDI 120  
Qy 121 IKGMSRPGVCSALNLCESLQKHLAELNHQKLESNKIPELDMTEVVAPFMANIPLLLY 180  
Db 121 IKGMSRPGVCSALNLCESLQKHLAELNHQKLESNKIPELDMTEVVAPFMANIPLLLY 180  
Qy 181 PQDGRSKPOPKONGDVCDICIQMVTDIQTAVRTNSTFVQALVEHVKCECDRLGPGMADI 240  
Db 181 PQDGRSKPOPKONGDVCDICIQMVTDIQTAVRTNSTFVQALVEHVKCECDRLGPGMADI 240  
Qy 241 CKNYISQYSEIATQMMHMQPKEICALVGFCDVEKEMPOTLVPAKVASKNVIIPALE 297  
Db 241 CKNYISQYSEIATQMMHMQPKEICALVGFCDVEKEMPOTLVPAKVASKNVIIPALE 300  
Qy 298 LVEPIKKEHVPKSDVYCEVCEFLVKEVTKLIDNNKTEKEILDAFKMCKSLPKSLSEEC 357  
Db 301 LVEPIKKEHVPKSDVYCEVCEFLVKEVTKLIDNNKTEKEILDAFKMCKSLPKSLSEEC 360  
Qy 358 QEVVDTYGSSTLSLLEEVSPBLVCSMLHLCSTGRLPALTVHVTQPKDGGFCEVCKKLVG 417  
Db 361 QEVVDTYGSSTLSLLEEVSPBLVCSMLHLCSTGRLPALTVHVTQPKDGGFCEVCKKLVG 420  
Qy 418 YLDRLNLEKNSTKQIBLALEKGCFLPDYQKQCDQFVAEYEPVLIIEILVEVMDPSFVCL 477  
Db 421 YLDRLNLEKNSTKQIBLALEKGCFLPDYQKQCDQFVAEYEPVLIIEILVEVMDPSFVCL 480  
Qy 478 KIGACPSAHPKLLGTEKICWGPSYWCQNTETAACNAVEHCKRHVWN 524  
Db 481 KIGACPSAHPKLLGTEKICWGPSYWCQNTETAACNAVEHCKRHVWN 527

RESULT 2  
A28716  
saposin precursor - rat  
N;Alternate names: cerebroside sulfate activator protein; co-beta-glucosidase; component ein (SAP); sphingolipid activator protein A2; sulfated glycoprotein 1; sulfatide sulfata  
N;Contains: prosaposin; saposin A; saposin B; saposin C; saposin D  
C;Species: Rattus norvegicus (Norway rat)  
C;Date: 10-Sep-1999 #sequence\_revision 10-Sep-1999 #text\_change 10-Sep-1999  
C;Accession: A28716  
R;Collard, M.W.; Sylvestre, S.R.; Tsuruta, J.K.; Griswold, M.D.  
Biochemistry 27, 4557-4564, 1988  
A;Title: Biosynthesis and molecular cloning of sulfated glycoprotein 1 secreted by rat s  
A;Reference number: A28716; MUID:89000647; PMID:3048385  
A;Accession: A28716  
A;Molecule type: mRNA  
A;Residues: 1-554 <COL>  
A;Cross-references: GB:M19936; NID:g206904; PIDN:AAA42136.1; PID:g206905  
A;Note: parts of this sequence, including the amino end of the mature protein, were dete  
C;Function:  
A;Description: saposins bind sphingolipids, form hydrophilic complexes and make them acc  
A;Pathway: sphingolipid catabolism  
A;Note: saposins A and C (SAP-2) activate hydrolysis of glucocerebroside by beta-glucosy  
A;Note: saposin B (SAP-1) activates hydrolysis of galactocerebroside sulfate by arylsulfi  
A;Note: saposin D activates hydrolysis of sphingomyelin by sphingomyelin phosphodiesterase  
C;Superfamily: saposin; saposin repeat homology

C;Keywords: alternative splicing; glycoprotein; lysoprotein; lysosomal storage disease; lysosome; sph  
F;1-16/Domain: signal sequence #status predicted <SIG>  
F;17-554/Product: prosaposin #status predicted <PRO>  
F;55-148/Domain: saposin repeat homology <SAP1>  
F;60-143/Product: saposin A #status predicted <SAPA>  
F;189-280/Domain: saposin repeat homology <SAP2>  
F;194-273/Product: saposin B #status predicted <SAB1>  
F;306-397/Domain: saposin repeat homology <SAP3>  
F;310-389/Product: saposin C #status predicted <SAPC>  
F;431-522/Domain: saposin repeat homology <SAP4>  
F;437-514/Product: saposin D #status predicted <SAPD>  
F;63-138,66-132,94-106,439-512,442-506,470-481/Disulfide bonds: #status predicted  
F;80,214,331,456/Binding site: carbohydrate (Asn) (covalent) #status predicted  
F;197-270,200-264,229-240,314-387,317-381,345-356/Disulfide bonds: #status predicted

Query Match 71.6%; Score 1996; DB 1; Length 554;  
Best Local Similarity 66.3%; Pred. No. 4.1e-124;  
Matches 368; Conservative 77; Mismatches 78; Indels 32; Gaps 3;

Qy 1 MYALFLASLLGAALGAPVLGLKECTRGSAVWCNQVKTASDCGAVKHCLQTVWNKPTVKS 60  
Db 1 MYALFLASLLGAALGAPVLGLKECTRGSAVWCNQVKTASDCGAVKHCLQTVWNKPTVKS 60  
Qy 61 LPCDICVDVTAAGDMLKDNATEEILVLEKTCDWLPKPNMSASCKEIVDSYLPVILDI 120  
Db 61 LPCDICVDVTAAGDMLKDNATEEILVLEKTCDWLPKPNMSASCKEIVDSYLPVILDI 120  
Qy 121 IKGMSRPGVCSALNLCESLQKHLAELNHQKLESNKIPELDMTEVVAPFMANIPLLLY 180  
Db 121 IKGMSRPGVCSALNLCESLQKHLAELNHQKLESNKIPELDMTEVVAPFMANIPLLLY 179  
Qy 181 PQDGRSKPOPKONGDVCDICIQMVTDIQTAVRTNSTFVQALVEHVKCECDRLGPGMADI 240  
Db 180 PQDGRSKPOPKONGDVCDICIQMVTDIQTAVRTNSTFVQALVEHVKCECDRLGPGVSDI 239  
Qy 241 CKNYISQYSEIATQMMHMQPKEICALVGFCDVEKEMPOTLVPAKVASKNVIIPALEIVE 300  
Db 240 CKNYISQYSEIATQMMHMQPKEICVNVGFCDEVKVPMRTLVPAFAIKNIPALELTD 299  
Qy 301 PIKKEHVPKSDVYCEVCEFLVKEVTKLIDNNKTEKEILDAFKMCKSLPKSLSEECQEV 360  
Db 300 PYQDVTQANVIFCQVQCLVMKRLSELIINNATEELIKGLSKACSLPAPATKRCQEV 359  
Qy 361 VDTYGSSTLSLLEEVSPBLVCSMLHLCSTGRLPALTVHVTQPKDGGFCEVCKKLVG 394  
Db 360 LVTFGPSLLDVLMEHVPNPLCGVISLCSANPNLVGLTLEQPAALVSAIPKEAPPKQPE 419  
Qy 395 -----ALTVHVTQPKDGGFCEVCKKLVGYYLDRLNLEKNSTKQIBLALEKGCFLPDYQK 449  
Db 420 EPKQSALRAHVPPQKNGGFCVCKKLVIYLEHNLKSTKEEILAALEKGCFLPDYQK 479  
Qy 450 QCQDFVAEYEPVLIIEILVEVMDPSFVCLKIGACPSAHPKLLGTEKICWGPSYWCQNTETA 509  
Db 480 QCQDFVAEYEPVLIIEILVEVMDPSFVCSKIGVCPKSAIKLLGLTGTEKCVMGVPGWCQNSETA 539  
Qy 510 AQCNVAHECHCKRHVWN 524  
Db 540 ARCNVADHCKRHVWN 554

RESULT 3  
JH0604  
saposin precursor - mouse  
N;Alternate names: cerebroside sulfate activator protein; co-beta-glucosidase; component ein (SAP); sphingolipid activator protein A2; sulfated glycoprotein 1; sulfatide sulfata  
N;Contains: prosaposin; saposin A; saposin B; saposin C; saposin D  
C;Species: Mus musculus (house mouse)  
C;Date: 10-Sep-1999 #sequence\_revision 10-Sep-1999 #text\_change 10-Sep-1999  
C;Accession: JH0604  
R;Tsuda, M.; Sakiyama, T.; Endo, H.; Kitagawa, T.  
Biochem. Biophys. Res. Commun. 184, 1266-1272, 1992  
A;Title: The primary structure of mouse saposin.  
A;Reference number: JH0604; MUID:92272718; PMID:1590788





A:Molecule type: protein  
A:Residues: 1-79 <STE>  
A:Experimental source: kidney  
A>Note: sequence extracted from NCBI backbone (NCBIP:129597)  
C:Superfamily: saposin; saposin repeat homology  
F:1-79/Domain: saposin repeat homology <SAP>

Query Match 13.9%; Score 389; DB 2; Length 79;  
Best Local Similarity 88.6%; Pred. No. 3.6e-19;  
Matches 70; Conservative 5; Mismatches 4; Indels 0; Gaps 0;

Qy 195 GDVQCDCIQWTDITQAVRTNSTFVQALVEHVKECDRLGPGMADICKNVIQSYSEIAIQ 254  
Db 1 GDVQCDCIQWTDITQAVRTNSTFVQALVEHVKECDRLGPGMADICKNVIQSYSEIAIQ 60  
Qy 255 MMEHMQPKKEICALVGFCDE 273  
Db 61 MMEHMQPKDICGLVGFCEE 79

RESULT 6  
S02766  
A:Molecule type: protein  
A:Residues: 1-376 <EMR>  
A:Cross-references: EMBL:X14778; NID:957284; PIDN:CAA32885.1; PID:g57285  
C:Superfamily: pulmonary surfactant protein B; saposin repeat homology  
F:1-19/Domain: signal sequence #status predicted <SIG>  
F:20-190/Domain: propeptide #status predicted <PRO>  
F:59-151/Domain: saposin repeat homology <SAP1>  
F:190-277/Domain: saposin repeat homology <SAP2>  
F:191-269/Product: pulmonary surfactant protein B #status predicted <MAT>  
F:286-371/Domain: saposin repeat homology <SAP3>

Query Match 13.6%; Score 379; DB 2; Length 376;  
Best Local Similarity 22.6%; Pred. No. 1.2e-17;  
Matches 116; Conservative 75; Mismatches 158; Indels 164; Gaps 16;

Qy 4 LFLILASLGAALAGPVLGLKCTGSAVQNVKTSDCGAVKICLOTWNKPTVKSLPC 63  
Db 9 LLLLPTLCSLGAATESASSPDCAQSPKFCQSLQEQAIQCRALGHCLQEVWGHAGANDL-C 67  
Qy 64 DICQDVVTAAGDMLKDKNATEEILVLEKTCQDLKPKNMSASCKEIVDSYLPVLDLIIG 123  
Db 68 QECEDIHLLTKYTKEDAFQDTIRKFLQECEDILPKLLVPRCQVLDVLPVLDIVFQG 127  
Qy 124 EMSRPGVCSALNICESLQKHLAELNHQKLESNKPELDMTEYVAPFMANIPILLYPOD 183  
Db 128 QI-KPKAICSHVGLC-----PL--GQT 146  
Qy 184 GPRSKPQKQNDGVQCDCIQWTDITQAVRTNSTFVQALVEHVKECDRLGPGMADICKN 243  
Db 147 KPEQKPE----- 153  
Qy 244 YISQSEIAIQMMHMQPKKEICALVGFCDEVKEMPMQTLVPAKVASKNVIPALB---LVE 300  
Db 154 -----MLDAIPNPLNKLKLVLPALPGAFAR 178  
Qy 301 P-----IKKHEVPAKSDVYCEVEFLVKEVTKLIDNNKTEKILD-AFOKMCCKLPKSL 353  
Db 179 PGPHQTDLSEQLPIPLP-FCWLKRTLKRVQAVI-----PKGVLAVALVSQVCHVPLV 232  
Qy 354 SEECQEVVDYTGSSILLLEVEVSPVCSMLHLC-----GTRLPALTWHVTO-PKGG 407  
Db 233 GGICQCLAERYTVLLDALLGRVVPQVLCGLVLRCSCTADAIQPALPALEPLIEKWPLODT 292

Qy 408 FCEVCKKLVGYLDRLNLEKNSTKQEIILAALEKGC-SFLPDPYQKQCDQVAYEYEVLTIEL 466  
Db 293 ECHFKCKSVI-----NQAWNTSEQAMPQAMHQAQLRFLWD--RQKCEQFVQHMPQLLALV 345  
Qy 467 VEVDPSFVCLKIGACPSAHKPLLGTGKCIWGP 499  
Db 346 PRSQDAHTSCQALGVCEAPASPL----QCPQTP 374

RESULT 7  
LNHUB  
A:Molecule type: protein  
A:Residues: 1-381 <PIL>  
A:Cross-references: GB:M24461  
A>Note: the codon given for residue 131 (ATT) is inconsistent with the authors' translation  
A>Note: this protein is encoded by a single gene  
R:Jacobs, K.A.; Phelps, D.S.; Steinbrink, R.; Fisch, J.; Kriz, J.; Mitsock, L.; Dougherty, J. Biol. Chem. 262, 9808-9811, 1987  
A:Title: Isolation of a cDNA clone encoding a high molecular weight precursor to a 6-kDa  
A:Reference number: A28461; MUID:87250653; PMID:3597440  
A:Accession: A28461  
A:Molecule type: mRNA  
A:Residues: 1-227, 'A', 229-381 <JAC>  
A:Cross-references: GB:J02761; NID:G190673; PIDN:AAA60212.1; PID:g190674  
A>Note: part of this sequence, including the amino end of the mature protein, was confirmed by R:Glasser, S.W.; Korfhagen, T.R.; Weaver, T.; Pilot-Matias, T.; Fox, J.L.; Whitsett, J.A. Proc. Natl. Acad. Sci. U.S.A. 84, 4007-4011, 1987  
A:Title: cDNA and deduced amino acid sequence of human pulmonary surfactant-associated protein B  
A:Reference number: A27794; MUID:87231940; PMID:3035561  
A:Accession: A27794  
A:Molecule type: mRNA  
A:Residues: 'EPR', 99-317, 'L', 319-381 <GLA>  
A:Cross-references: GB:M16764; NID:G338410; PIDN:AAA88099.1; PID:g338411  
A>Note: 131-11e was also found  
A>Note: part of this sequence, including the amino end of the mature protein, was confirmed by R:Revak, S.D.; Merritt, T.A.; Degryse, E.; Stefani, L.; Courtney, M.; Hallman, M.; Cochran, J. Clin. Invest. 81, 826-833, 1988  
A:Title: Use of human surfactant low molecular weight apoproteins in the reconstitution of pulmonary surfactant  
A:Reference number: A27592; MUID:88139786; PMID:3343343  
A:Accession: A27592  
A:Molecule type: mRNA  
A:Residues: 139-177, 'V', 179-227, 'A', 228-381 <REV>  
A:Cross-references: GB:M19097  
A>Note: part of this sequence, including the amino end of the mature protein, was confirmed by R:Mizumoto, M.; Adachi, H.  
A:Title: Primary structure of a hydrophobic 6kDa apoprotein (SP6) of human pulmonary surfactant  
A:Reference number: JU0162  
A:Accession: JU0162  
A:Molecule type: protein  
A:Residues: 201-207, 'X', 209-210, 'X', 212-227, 'A', 229-234, 'X', 236-245, 'X', 247, 'L', 249-253, R:Johansson, J.; Joernvall, H.; Curstedt, T.  
FEBS Lett. 301, 165-167, 1992  
A:Title: Human surfactant polypeptide SP-B. Disulfide bridges, C-terminal end, and peptide  
A:Reference number: S21134; MUID:92233937; PMID:1568474  
A:Accession: S21134  
A>Status: preliminary  
A:Molecule type: protein  
A:Residues: 201-227, 'I', 229-279 <JOH>  
A>Note: 228-Ala was also found  
C:Comment: Pulmonary surfactant is a complex of phospholipids and proteins that lowers the





QY 244 YISQYSEIAIOMMMHQPKEICALVGFCEVDEKEMPMQTLVPAKASKNVIPALELVEPIK 303  
Db 176 -----HTQ-----DLSAQPFPIPLP-----190  
QY 304 KHEVPAKSDVYCEVCFVKEVTKLIDNNKTEKILDAFDKWCCKLPKSLSECEVVD 362  
Db 191 -----LWLCRTLLKRIQAMT-----PKGVLAQVAVQVCHVPLVVGIGICQCLAE 235  
QY 363 TYGSSILSILLEVSPVCSMLHLCS-----GTRLPALTIVHTQ---PRDGGFCVCKKL 415  
Db 236 RYTVILLEVLLGHVLPQVGLVLRCSVSDSIGOVPTLEALPGEWLPQDPE-CHLCMSV 294  
QY 416 VGYLDRLNLEKNSTKQBIILAA-----LEKGCSEFLPDYKQCDQFVAEYEPVLIIEIVEM 470  
Db 295 TTQA-RNISEQTPQAVHACLSQLDK-----OECEQFVAHAAP-----AA 335  
QY 471 DFSFVCLKTGACPSAHKPLLTGE-----KCIWGPSY 501  
Db 336 EPAVQGL-----GCP-----RNLPGPGRVATLSLQCIQSPHF 370

## RESULT 13

T46069

hypothetical protein T18N14.110 - Arabidopsis thaliana

C:Species: Arabidopsis thaliana (mouse-ear cress)

C&gt;Date: 04-Feb-2000 #sequence\_revision 04-Feb-2000 #text\_change 04-Feb-2000

C:Accession: T46069

R:Delseny, M.; Berger, C.; Cooke, R.; Grellert, F.; Laudie, M.; Mewes, H.W.; Lemcke, K.;

submitted to the Protein Sequence Database, December 1999

A:Reference number: Z23013

A:Accession: T46069

A&gt;Status: preliminary

A:Molecule type: DNA

A:Residues: 1-213 &lt;DEL&gt;

A:Cross-references: EMBL:AL132968

A:Experimental source: cultivar Columbia; BAC clone T18N14

C:Genetics:

A:Map position: 3

A:Introns: 31/1; 80/3; 146/3; 166/3

A:Note: T18N14.110

Query Match 7.6%; Score 213; DB 2; Length 213;

Best Local Similarity 29.7%; Pred. No. 5.1e-07;

Matches 57; Conservative 35; Mismatches 90; Indels 10; Gaps 6;

QY 292 VIPALELVEPIKHEVPAKSDVYCEVCFVKEVTKLIDNNKTEKILDAFDKWCCKLPK 351

Db 18 VSDARSFVDSITSEKVNKEDV-CTLCSEYVTDALSYLEKNVTVQABIIEIDLHRCSQL-R 75

QY 352 SLSECEQVVDYTGSSILSILLEVSPVCSMLHLCSGTRLPALTIVHTQPKDGGFCV 411

Db 76 GYSQQCISLVDDY-VPLFFQLSFSPHYFCRKNLGC--KVALVEEARQDS----CGV 128

QY 412 CKKLVGYLDRLNLEKNSTKQBIILAALEKGCSEFLPDYKQCDQFVAEYEPVLIIEIVEM 471

Db 129 CHRTVSEILIKLPDPTOLDIVELLKIGCKSLKN-YEKKCTLVFEYGLIILVNAEEFLV 187

QY 472 PSFVCLKIGACP 483

Db 188 KNDVCTLLRACP 199

## RESULT 14

T48201

hypothetical protein T20L15.70 - Arabidopsis thaliana

C:Species: Arabidopsis thaliana (mouse-ear cress)

C&gt;Date: 20-Apr-2000 #sequence\_revision 20-Apr-2000 #text\_change 20-Apr-2000

C:Accession: T48201

R:Bevan, M.; Peters, S.A.; van Staveren, M.; Dirkee, W.; Stiekema, W.; Bancroft, I.; Mew

submitted to the Protein Sequence Database, March 2000

A:Reference number: Z24488

A:Accession: T48201

A>Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-217 <BEV>  
A:Cross-references: EMBL:AL162351  
A:Experimental source: cultivar Columbia; BAC clone T20L15  
C:Genetics:  
A:Map position: 5  
A:Introns: 30/1; 79/3; 146/3; 166/3  
A:Note: T20L15.70

## Query Match

Best Local Similarity 7.3%; Score 203.5; DB 2; Length 217;

Matches 52; Conservative 40; Mismatches 82; Indels 13; Gaps 7;

QY 298 LVEPIKK-HEVPAKSDVYCEVCFVKEVTKLIDNNKTEKILDAFDKWCCKLPKSLSE 356

Db 25 LLEPFESAHQ-----DNQVCELCDDKYVTVLVIDYLDQYDNQNELVEALHISQIP-PLKKQ 79

QY 357 CQEVVDYTGSSILSILLEVSPVCSMLHLCSGTRLPALTIVHTQPKDGGFCVCKLV 416

Db 80 CLSMVDHY-TQLFFTOVSTIKSDQICKRLNLQAV-TPAFASQVHQ-----GNCEACRETV 133

QY 417 GYLDRLNLEKNSTKQBIILAALEKGCSEFLPDYKQCDQFVAEYEPVLIIEIVEMDPSPFC 476

Db 134 SEVTVKLPDPETKLIIRLLLECKECSL-NNYQDKCKGVFEYGLMLTDLQKFLKKDVC 192

QY 477 LKIGACP 483

Db 193 TILHVCP 199

## RESULT 15

T15677

hypothetical protein C28C12.5 - Caenorhabditis elegans

C:Species: Caenorhabditis elegans

C&gt;Date: 20-Sep-1999 #sequence\_revision 20-Sep-1999 #text\_change 20-Sep-1999

C:Accession: T15677

R:Miller, N.

submitted to the EMBL Data Library, November 1995

A:Description: The sequence of C. elegans cosmid C28C12.

A:Reference number: Z18387

A:Accession: T15677

A&gt;Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-402 &lt;ML&gt;

A:Cross-references: EMBL:U40797; NID:g1065916; PID:g1065921; PIDN:AAB37548.1; GSPDB:GN001

A:Experimental source: strain Bristol N2; clone C28C12

C:Genetics:

A:Gene: CESP:C28C12.5

A:Map position: 4

A:Introns: 29/3; 82/3; 124/3; 151/3; 258/3; 318/1; 372/3; 400/3

## Query Match

Best Local Similarity 6.3%; Score 176.5; DB 2; Length 402;

Matches 75; Conservative 58; Mismatches 156; Indels 81; Gaps 15;

QY 184 GPRSKPQKNGDVQCDCIQMTVDIQTAVTNTSTFVQALVEHVKEECDRLPGMGADICKN 243

Db 16 GAQSAPSP-----CESCKSNVQNFIDASKDRMQMAQLKYS-LSMLC--VGTSHQSDCSK 66

QY 244 YISQYSEIAIOMMMHMQPKEICALVGFCEVDEKEMPMQTLVPAKASKNVIPALELVEPI 302

Db 67 TLQKLDPIAYKLAPYLADTSVCSKLMQCGESQFSPRLARLAMLKXSEIVA---NDNI 133

QY 303 KHEVPAKSDVYCEVCFVKEVTKLIDNNKTEKILDAFDKWCCKLPKSLSECEVVD 362

Db 124 MRQEV-----CDECOASTAQIGKLVGDEFTTYAVKSTLQRFVCKSAGKAHKAC---N 172

QY 363 TYGSSILSILLEV-----SPELVCSMLHLCSGTRLPALTIVHTQPKDGGFCVCKLVG 417

Db 173 IFVSSVIPDLMTKMDWFTKELMCSNMGLCSATSKPAAREAPKQPAS-----EMWKSMM-- 226

QY 418 YLDRLNLEKNSTKQBIILAALEKGCSEFLPDYQKQ-----

```
Db      227  ---GMVKTSGBELMSCFE--CTLSADALLQEFIDKRGQTADDIQTVACNKMVANWTDG 280
Qy      451  CDQFVAEYEPVLIIEILVEVMDPSFVCLKIGACPSAHKPLLG---TEKCIWGPSYWCNTE 507
Db      281  CNDFVHMYSVLFITYNQFDGRGICTTMMHSCEKKENALVEMAMSEKVM/LG-----CEN-- 334
Qy      508  TAAQCNAVEH 517
Db      335  ---CKAVEH 340
```

Search completed: May 5, 2004, 13:34:33  
Job time : 21.9003 secs

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

## OM protein - protein search, using sw model

Run on: May 5, 2004, 13:14:53 ; Search time 55.7341 Seconds

(without alignments)  
2606.077 Million cell updates/sec

Title: US-09-743-684a-1

Perfect score: 2789  
Sequence: 1 MYALFLASLLGALAGVLT.....NTETRAQCNAYEHCKRHVMN 524

## Scoring table:

BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 1138120 seqs, 277189581 residues

Total number of hits satisfying chosen parameters: 1138120

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :	Published Applications AA:*
1:	/cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep:*
2:	/cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pep:*
3:	/cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep:*
4:	/cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep:*
5:	/cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep:*
6:	/cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB.pep:*
7:	/cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep:*
8:	/cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep:*
9:	/cgn2_6/ptodata/2/pubpaa/US09_PUBCOMB.pep:*
10:	/cgn2_6/ptodata/2/pubpaa/US09B_PUBCOMB.pep:*
11:	/cgn2_6/ptodata/2/pubpaa/US09C_PUBCOMB.pep:*
12:	/cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pep:*
13:	/cgn2_6/ptodata/2/pubpaa/US10_PUBCOMB.pep:*
14:	/cgn2_6/ptodata/2/pubpaa/US10B_PUBCOMB.pep:*
15:	/cgn2_6/ptodata/2/pubpaa/US10C_PUBCOMB.pep:*
16:	/cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pep:*
17:	/cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep:*
18:	/cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2789	100.0	524	9	US-09-870-759-60 Sequence 60, Appl
2	2789	100.0	524	10	US-09-751-708A-60 Sequence 386, App
3	2789	100.0	524	12	US-10-267-502-386 Sequence 61, Appl
4	2777.5	99.6	527	9	US-09-870-759-61 Sequence 61, Appl
5	2777.5	99.6	527	10	US-09-751-708A-61 Sequence 61, Appl
6	2777.5	99.6	527	14	US-10-060-036-73 Sequence 73, Appl
7	2757.5	98.9	523	9	US-09-767-007A-2 Sequence 40, Appl
8	2449.5	87.8	479	10	US-09-978-418-40 Sequence 176, App
9	1969	70.6	554	14	US-10-205-194-176 Sequence 1, Appl
10	1201	43.1	521	15	US-10-276-162-1 Sequence 903, App
11	1139.5	40.9	531	11	US-09-833-245-903 Sequence 340, App
12	1093	39.2	209	14	US-10-043-487-340 Sequence 8, Appl
13	834.5	29.2	362	12	US-10-332-426-8 Sequence 385, App
14	552	19.8	953	12	US-10-267-502-385 Sequence 4529, Ap
15	545	19.5	241	15	US-10-108-260A-4529

16	539.5	19.3	156	12	US-09-925-298-644 Sequence 644, App
17	539.5	19.3	156	14	US-10-102-806-644 Sequence 644, App
18	412	14.8	80	9	US-09-767-007A-3 Sequence 3, Appl
19	412	14.8	80	9	US-09-753-126-3 Sequence 3, Appl
20	412	14.8	80	15	US-10-330-697-3 Sequence 3, Appl
21	412	14.8	592	9	US-09-753-126-4 Sequence 4, Appl
22	412	14.8	592	15	US-10-330-697-4 Sequence 4, Appl
23	349.5	12.5	381	15	US-10-236-031B-10 Sequence 10, Appl
24	349.5	12.5	381	15	US-10-295-027-1223 Sequence 1223, Ap
25	235.5	8.4	216	12	US-10-424-599-276328 Sequence 276328, Ap
26	211	7.6	40	10	US-09-780-438A-1 Sequence 1, Appl
27	211	7.6	212	12	US-10-424-599-157904 Sequence 157904, A
28	204.5	7.3	246	12	US-10-425-114-41077 Sequence 41077, A
29	201	7.2	38	10	US-09-780-438A-2 Sequence 2, Appl
30	201	7.2	212	12	US-10-424-599-157906 Sequence 157906, A
31	197	7.1	252	12	US-10-425-114-55606 Sequence 55606, A
32	184	6.6	243	12	US-10-425-114-67929 Sequence 67929, A
33	175.5	6.3	265	12	US-10-425-114-51535 Sequence 51535, A
34	169.5	6.1	236	12	US-10-424-599-227975 Sequence 227975, A
35	168.5	6.0	181	8	US-08-488-123-12 Sequence 12, Appl
36	165.5	5.9	246	12	US-10-425-114-60351 Sequence 60351, A
37	163	5.8	514	14	US-10-339-351-3 Sequence 3, Appl
38	162	5.8	514	14	US-10-339-351-1 Sequence 1, Appl
39	159.5	5.7	281	12	US-10-424-599-205602 Sequence 205602, A
40	157.5	5.6	282	12	US-10-424-599-205603 Sequence 205603, A
41	156.5	5.6	507	12	US-10-424-599-245585 Sequence 245585, A
42	152	5.4	471	12	US-10-425-114-71427 Sequence 71427, A
43	151	5.4	393	12	US-10-425-114-65142 Sequence 65142, A
44	147.5	5.3	223	12	US-10-425-114-40454 Sequence 40454, A
45	146.5	5.3	295	12	US-10-425-114-57134 Sequence 57134, A

## ALIGNMENTS

RESULT 1	
US-09-870-759-60	
; Sequence 60, Application US/09870759	
; Patient No. US20020177551A1	
; GENERAL INFORMATION:	
; APPLICANT: TERMAN, David S	
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATMENT OF NEOPLASTIC DISEASE	
; FILE REFERENCE: 870759	
; CURRENT APPLICATION NUMBER: US/09/870,759	
; CURRENT FILING DATE: 2002-01-14	
; PRIOR APPLICATION NUMBER: US 60/208,128	
; PRIOR FILING DATE: 2000-05-30	
; NUMBER OF SEQ ID NOS: 166	
; SOFTWARE: PatentIn version 3.1	
; SEQ ID NO 60	
; LENGTH: 524	
; TYPE: PRT	
; ORGANISM: Homo sapiens	
US-09-870-759-60	
Query Match	100.0%; Score 2789; DB 9; Length 524;
Best Local Similarity	100.0%; Pred. No. 3.6e-239;
Matches 524; Conservative 0; Mismatches 0; Indels 0; Gaps 0;	
OY	1 MYALFLASLLGALAGVLTGKECTRGSAWCVKTAPOCAVKGICLQTVWKNPKPYKS 60
DB	1 MYALFLASLLGALAGVLTGKECTRGSAWCVKTAPOCAVKGICLQTVWKNPKPYKS 60
OY	61 LPDCICKDVTAAGDMLKDNATEEILVLYLEKTCQDWLKPNNASCKEIVDSYLVILDI 120
DB	61 LPDCICKDVTAAGDMLKDNATEEILVLYLEKTCQDWLKPNNASCKEIVDSYLVILDI 120
OY	121 IKGMSRSGEVSALNLCESLQKHLAEINHOXKLESNNKIPELDMTEVAPPMANIPILLY 180
DB	121 IKGMSRSGEVSALNLCESLQKHLAEINHOXKLESNNKIPELDMTEVAPPMANIPILLY 180
OY	181 FODGRSPKPKNDGVCODICQWTDIQTAVRTNSTFVQALVBEHVEKCRLEGGMADI 240
DB	181 FODGRSPKPKNDGVCODICQWTDIQTAVRTNSTFVQALVBEHVEKCRLEGGMADI 240

Db 181 PDGPRSKPQPKDNGVCDQCIQWMTDIOGAVRNTSTFVQALVEHYKECDRLGPGMADI 240  
Qy 241 CKNYISQYSEIAIQMMHMQPKETCALVGFCEVKEKMPQTLVPARVASKNVLPALAEVE 300  
Db 241 CKNYISQYSEIAIQMMHMQPKETCALVGFCEVKEKMPQTLVPARVASKNVLPALAEVE 300  
Qy 301 PIKKEHPAKSDVYCEVCEFLVKEVTKLIDNNKTEKEIIDAFAFKMCSKLPKSLSECCOV 360  
Db 301 PIKKEHPAKSDVYCEVCEFLVKEVTKLIDNNKTEKEIIDAFAFKMCSKLPKSLSECCOV 360  
Qy 361 VDTYSSILSLLEBVSPELVCSMLHCSGTRLPALTVAHTQPKDGFCEVCKKLVGYLD 420  
Db 361 VDTYSSILSLLEBVSPELVCSMLHCSGTRLPALTVAHTQPKDGFCEVCKKLVGYLD 420  
Qy 421 RNLEKSTKOEITLAEKSGSFLPDYQKQCDQFAVEYEPVLIELIVEMDPSFVCLKIG 480  
Db 421 RNLEKSTKOEITLAEKSGSFLPDYQKQCDQFAVEYEPVLIELIVEMDPSFVCLKIG 480  
Qy 481 ACPSAHKPILGTEKCIWGPSYWCQNTETAACNAVEHCKRHVWN 524  
Db 481 ACPSAHKPILGTEKCIWGPSYWCQNTETAACNAVEHCKRHVWN 524

## RESULT 2

US-09-751-708A-60  
; Sequence 60, Application US/09751708A  
; Publication No. US20030157113A1  
; GENERAL INFORMATION:  
; APPLICANT: TERMAN, David S  
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATMENT OF NEOPLASTIC DISEASE  
; FILE REFERENCE: 751708  
; CURRENT APPLICATION NUMBER: US/09/751,708A  
; CURRENT FILING DATE: 2002-10-15  
; PRIOR APPLICATION NUMBER: US 60/173,371  
; PRIOR FILING DATE: 1999-12-28  
; NUMBER OF SEQ ID NOS: 166  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO: 60  
; LENGTH: 524  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-751-708A-60

Query Match 100.0%; Score 2789; DB 10; Length 524;  
Best Local Similarity 100.0%; Pred. No. 3,66-239;  
Matches 524; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 1 MYALFLASLIGALAGPVLGLKECTRGSAVWCQNVKTAASDCGAVKHCLQTVWNNKPTVKS 60  
Db 1 MYALFLASLIGALAGPVLGLKECTRGSAVWCQNVKTAASDCGAVKHCLQTVWNNKPTVKS 60  
Qy 61 LPDICKDVVTAAGDMLKDNATEEBILVYLEKTCDWLPKPNMSASCKEIVDSYLPVILDI 120  
Db 61 LPDICKDVVTAAGDMLKDNATEEBILVYLEKTCDWLPKPNMSASCKEIVDSYLPVILDI 120  
Qy 121 IKGEMSRPGEVCSALNLCESLQKHIAELNHQKQLESNKIPBLDMTEVVAPEMANIPILLY 180  
Db 121 IKGEMSRPGEVCSALNLCESLQKHIAELNHQKQLESNKIPBLDMTEVVAPEMANIPILLY 180  
Qy 181 PDGPRSKPQPKDNGVCDQCIQWMTDIOGAVRNTSTFVQALVEHYKECDRLGPGMADI 240  
Db 181 PDGPRSKPQPKDNGVCDQCIQWMTDIOGAVRNTSTFVQALVEHYKECDRLGPGMADI 240  
Qy 241 CKNYISQYSEIAIQMMHMQPKETCALVGFCEVKEKMPQTLVPARVASKNVLPALAEVE 300  
Db 241 CKNYISQYSEIAIQMMHMQPKETCALVGFCEVKEKMPQTLVPARVASKNVLPALAEVE 300  
Qy 301 PIKKEHPAKSDVYCEVCEFLVKEVTKLIDNNKTEKEIIDAFAFKMCSKLPKSLSECCOV 360  
Db 301 PIKKEHPAKSDVYCEVCEFLVKEVTKLIDNNKTEKEIIDAFAFKMCSKLPKSLSECCOV 360  
Qy 361 VDTYSSILSLLEBVSPELVCSMLHCSGTRLPALTVAHTQPKDGFCEVCKKLVGYLD 420  
Db 361 VDTYSSILSLLEBVSPELVCSMLHCSGTRLPALTVAHTQPKDGFCEVCKKLVGYLD 420

Db 361 VDTYSSILSLLEBVSPELVCSMLHCSGTRLPALTVAHTQPKDGFCEVCKKLVGYLD 420  
Qy 421 RNLEKSTKOEITLAEKSGSFLPDYQKQCDQFAVEYEPVLIELIVEMDPSFVCLKIG 480  
Db 421 RNLEKSTKOEITLAEKSGSFLPDYQKQCDQFAVEYEPVLIELIVEMDPSFVCLKIG 480  
Qy 481 ACPSAHKPILGTEKCIWGPSYWCQNTETAACNAVEHCKRHVWN 524  
Db 481 ACPSAHKPILGTEKCIWGPSYWCQNTETAACNAVEHCKRHVWN 524

## RESULT 3

US-10-267-502-386  
; Sequence 386, Application US/10267502  
; Publication No. US20040071700A1  
; GENERAL INFORMATION:  
; APPLICANT: Kim, Jaeseob  
; TITLE OF INVENTION: Obesity Linked Genes  
; FILE REFERENCE: LSD-07416  
; CURRENT APPLICATION NUMBER: US/10/267,502  
; CURRENT FILING DATE: 2003-01-27  
; NUMBER OF SEQ ID NOS: 439  
; SOFTWARE: PatentIn version 3.2  
; SEQ ID NO: 386  
; LENGTH: 524  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-267-502-386

Query Match 100.0%; Score 2789; DB 12; Length 524;  
Best Local Similarity 100.0%; Pred. No. 3,66-239;  
Matches 524; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MYALFLASLIGALAGPVLGLKECTRGSAVWCQNVKTAASDCGAVKHCLQTVWNNKPTVKS 60  
Db 1 MYALFLASLIGALAGPVLGLKECTRGSAVWCQNVKTAASDCGAVKHCLQTVWNNKPTVKS 60  
Qy 61 LPDICKDVVTAAGDMLKDNATEEBILVYLEKTCDWLPKPNMSASCKEIVDSYLPVILDI 120  
Db 61 LPDICKDVVTAAGDMLKDNATEEBILVYLEKTCDWLPKPNMSASCKEIVDSYLPVILDI 120  
Qy 121 IKGEMSRPGEVCSALNLCESLQKHIAELNHQKQLESNKIPBLDMTEVVAPEMANIPILLY 180  
Db 121 IKGEMSRPGEVCSALNLCESLQKHIAELNHQKQLESNKIPBLDMTEVVAPEMANIPILLY 180  
Qy 181 PDGPRSKPQPKDNGVCDQCIQWMTDIOGAVRNTSTFVQALVEHYKECDRLGPGMADI 240  
Db 181 PDGPRSKPQPKDNGVCDQCIQWMTDIOGAVRNTSTFVQALVEHYKECDRLGPGMADI 240  
Qy 241 CKNYISQYSEIAIQMMHMQPKETCALVGFCEVKEKMPQTLVPARVASKNVLPALAEVE 300  
Db 241 CKNYISQYSEIAIQMMHMQPKETCALVGFCEVKEKMPQTLVPARVASKNVLPALAEVE 300  
Qy 301 PIKKEHPAKSDVYCEVCEFLVKEVTKLIDNNKTEKEIIDAFAFKMCSKLPKSLSECCOV 360  
Db 301 PIKKEHPAKSDVYCEVCEFLVKEVTKLIDNNKTEKEIIDAFAFKMCSKLPKSLSECCOV 360  
Qy 361 VDTYSSILSLLEBVSPELVCSMLHCSGTRLPALTVAHTQPKDGFCEVCKKLVGYLD 420  
Db 361 VDTYSSILSLLEBVSPELVCSMLHCSGTRLPALTVAHTQPKDGFCEVCKKLVGYLD 420  
Qy 421 RNLEKSTKOEITLAEKSGSFLPDYQKQCDQFAVEYEPVLIELIVEMDPSFVCLKIG 480  
Db 421 RNLEKSTKOEITLAEKSGSFLPDYQKQCDQFAVEYEPVLIELIVEMDPSFVCLKIG 480  
Qy 481 ACPSAHKPILGTEKCIWGPSYWCQNTETAACNAVEHCKRHVWN 524  
Db 481 ACPSAHKPILGTEKCIWGPSYWCQNTETAACNAVEHCKRHVWN 524

RESULT 4  
US-09-870-759-61



```
; Sequence 61, Application US/09870759
; Patent No. US20020177551A1
; GENERAL INFORMATION:
; APPLICANT: TERMAN, David S
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATMENT OF NEOPLASTIC DISEASE
; FILE REFERENCE: 870759
; CURRENT APPLICATION NUMBER: US/09/870,759
; CURRENT FILING DATE: 2002-01-14
; PRIOR APPLICATION NUMBER: US 60/208,128
; PRIOR FILING DATE: 2000-05-30
; NUMBER OF SEQ ID NOS: 166
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 61
; LENGTH: 527
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-09-870-759-61

Query Match          99.4%; Score 2777.5; DB 9; Length 527;
Best Local Similarity 99.4%; Pred. No. 3.9e-238;
Matches 524; Conservative 0; Mismatches 0; Indels 3; Gaps 1;

QY 1 MYALFLASLIGALAGVTLGKECTRGSAVWCQNVKTASDCGAVKGLQTVNKPVTYS 60
DB 1 MYALFLASLIGALAGVTLGKECTRGSAVWCQNVKTASDCGAVKGLQTVNKPVTYS 60
QY 61 LPDICKVVTAAAGDMLKDNATEBEILVYLEKTCMLPKPMNASCKEIVSYLPVILI 120
DB 61 LPDICKVVTAAAGDMLKDNATEBEILVYLEKTCMLPKPMNASCKEIVSYLPVILI 120
QY 121 IKGEMSRGVECSALNLCESLQKHLAEINHOKELESNKIPBLDMTEVVAAPPMANIPLLY 180
DB 121 IKGEMSRGVECSALNLCESLQKHLAEINHOKELESNKIPBLDMTEVVAAPPMANIPLLY 180
QY 181 PODGRSRKPOPKNDGVCDQCIQWVTDIQTAVRTNSTFVQALVEHVKECDRLGPGMADI 240
DB 181 PODGRSRKPOPKNDGVCDQCIQWVTDIQTAVRTNSTFVQALVEHVKECDRLGPGMADI 240
QY 241 CKNYISQYSEIAIQMMMHMODQOPKEICALVGFCDVEKEMQOTLVPAKVASKNVIIPALE 297
DB 241 CKNYISQYSEIAIQMMMHMODQOPKEICALVGFCDVEKEMQOTLVPAKVASKNVIIPALE 297
QY 298 LVEPIKKEHVPAKSDVYCEVCEFLVKEVTKLIDNNKTEKEILDADFDDKSKLPSLSSEC 357
DB 298 LVEPIKKEHVPAKSDVYCEVCEFLVKEVTKLIDNNKTEKEILDADFDDKSKLPSLSSEC 357
QY 361 QEVVDTYSSILSLILEVSPELVCSMLHLCSGTRLPALTVHTQPKDGFCEVCKLVG 420
DB 361 QEVVDTYSSILSLILEVSPELVCSMLHLCSGTRLPALTVHTQPKDGFCEVCKLVG 420
QY 417 358 QEVVDTYSSILSLILEVSPELVCSMLHLCSGTRLPALTVHTQPKDGFCEVCKLVG 417
DB 417 358 QEVVDTYSSILSLILEVSPELVCSMLHLCSGTRLPALTVHTQPKDGFCEVCKLVG 417
QY 477 418 YIDRLNENKSTKQETILAEKGCSPFLPDYQKCDQFVAEYEPVILIEILVEMDPSFVCL 477
DB 477 418 YIDRLNENKSTKQETILAEKGCSPFLPDYQKCDQFVAEYEPVILIEILVEMDPSFVCL 477
QY 480 421 YIDRLNENKSTKQETILAEKGCSPFLPDYQKCDQFVAEYEPVILIEILVEMDPSFVCL 480
DB 480 421 YIDRLNENKSTKQETILAEKGCSPFLPDYQKCDQFVAEYEPVILIEILVEMDPSFVCL 480
QY 524 478 KIGACPSAHKPLGTGTEKCIWGPSYWCQNTETAAQCNAVEHCKRHWN 524
DB 524 478 KIGACPSAHKPLGTGTEKCIWGPSYWCQNTETAAQCNAVEHCKRHWN 524
; RESULT 5
; US-09-751-708A-61
; Sequence 61, Application US/09751708A
; Publication No. US20030157113A1
; GENERAL INFORMATION:
; APPLICANT: TERMAN, David S
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATMENT OF NEOPLASTIC DISEASE
; FILE REFERENCE: 751708
; CURRENT APPLICATION NUMBER: US/09/751,708A
; CURRENT FILING DATE: 2002-10-15
; PRIOR APPLICATION NUMBER: US 60/173,371
; PRIOR FILING DATE: 1999-12-28
; NUMBER OF SEQ ID NOS: 166
; SOFTWARE: PatentIn version 3.1
```

```
; SEQ ID NO 61
; LENGTH: 527
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-09-751-708A-61

Query Match          99.6%; Score 2777.5; DB 10; Length 527;
Best Local Similarity 99.4%; Pred. No. 3.9e-238;
Matches 524; Conservative 0; Mismatches 0; Indels 3; Gaps 1;

QY 1 MYALFLASLIGALAGVTLGKECTRGSAVWCQNVKTASDCGAVKGLQTVNKPVTYS 60
DB 1 MYALFLASLIGALAGVTLGKECTRGSAVWCQNVKTASDCGAVKGLQTVNKPVTYS 60
QY 61 LPDICKVVTAAAGDMLKDNATEBEILVYLEKTCMLPKPMNASCKEIVSYLPVILI 120
DB 61 LPDICKVVTAAAGDMLKDNATEBEILVYLEKTCMLPKPMNASCKEIVSYLPVILI 120
QY 121 IKGEMSRGVECSALNLCESLQKHLAEINHOKELESNKIPBLDMTEVVAAPPMANIPLLY 180
DB 121 IKGEMSRGVECSALNLCESLQKHLAEINHOKELESNKIPBLDMTEVVAAPPMANIPLLY 180
QY 181 PODGRSRKPOPKNDGVCDQCIQWVTDIQTAVRTNSTFVQALVEHVKECDRLGPGMADI 240
DB 181 PODGRSRKPOPKNDGVCDQCIQWVTDIQTAVRTNSTFVQALVEHVKECDRLGPGMADI 240
QY 241 CKNYISQYSEIAIQMMMHMODQOPKEICALVGFCDVEKEMQOTLVPAKVASKNVIIPALE 297
DB 241 CKNYISQYSEIAIQMMMHMODQOPKEICALVGFCDVEKEMQOTLVPAKVASKNVIIPALE 297
QY 298 LVEPIKKEHVPAKSDVYCEVCEFLVKEVTKLIDNNKTEKEILDADFDDKSKLPSLSSEC 357
DB 298 LVEPIKKEHVPAKSDVYCEVCEFLVKEVTKLIDNNKTEKEILDADFDDKSKLPSLSSEC 357
QY 361 QEVVDTYSSILSLILEVSPELVCSMLHLCSGTRLPALTVHTQPKDGFCEVCKLVG 420
DB 361 QEVVDTYSSILSLILEVSPELVCSMLHLCSGTRLPALTVHTQPKDGFCEVCKLVG 420
QY 417 358 QEVVDTYSSILSLILEVSPELVCSMLHLCSGTRLPALTVHTQPKDGFCEVCKLVG 417
DB 417 358 QEVVDTYSSILSLILEVSPELVCSMLHLCSGTRLPALTVHTQPKDGFCEVCKLVG 417
QY 477 418 YIDRLNENKSTKQETILAEKGCSPFLPDYQKCDQFVAEYEPVILIEILVEMDPSFVCL 477
DB 477 418 YIDRLNENKSTKQETILAEKGCSPFLPDYQKCDQFVAEYEPVILIEILVEMDPSFVCL 477
QY 480 421 YIDRLNENKSTKQETILAEKGCSPFLPDYQKCDQFVAEYEPVILIEILVEMDPSFVCL 480
DB 480 421 YIDRLNENKSTKQETILAEKGCSPFLPDYQKCDQFVAEYEPVILIEILVEMDPSFVCL 480
QY 524 478 KIGACPSAHKPLGTGTEKCIWGPSYWCQNTETAAQCNAVEHCKRHWN 524
DB 524 478 KIGACPSAHKPLGTGTEKCIWGPSYWCQNTETAAQCNAVEHCKRHWN 524
; RESULT 6
; US-10-060-036-73
; Sequence 73, Application US/10060036
; Publication No. US20030073144A1
; GENERAL INFORMATION:
; APPLICANT: Benson, Darin R.
; APPLICANT: Kalos, Michael D.
; APPLICANT: Lodes, Michael J.
; APPLICANT: Persing, David H.
; APPLICANT: Hepler, William T.
; APPLICANT: Jiang, Yugu
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; FILE REFERENCE: 210121.566
; CURRENT APPLICATION NUMBER: US/10/060,036
; CURRENT FILING DATE: 2002-01-30
; NUMBER OF SEQ ID NOS: 4560
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 73
; LENGTH: 527
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-10-060-036-73

Query Match          99.6%; Score 2777.5; DB 14; Length 527;
Best Local Similarity 99.4%; Pred. No. 3.9e-238;
```

Matches 524; Conservative 0; Mismatches 0; Indels 3; Gaps 1;

QY 1 MYALFLASLILGALAGVTLGLKECTRGSAAVWCQNVKTASDCGAVKHCLQTVMNKPITYKS 60  
 Db 1 MYALFLASLILGALAGVTLGLKECTRGSAAVWCQNVKTASDCGAVKHCLQTVMNKPITYKS 60

QY 61 LPDCICKDVVTAAGMDLKDNAATEEELIVLEKTCMDLPRKPMNSASCKEIVDSYLPVILDI 120  
 Db 61 LPDCICKDVVTAAGMDLKDNAATEEELIVLEKTCMDLPRKPMNSASCKEIVDSYLPVILDI 120

QY 121 IKGEMSRPGEVCSALNLCESLQKHLAELNHOKLESNKIPBLDMEVVAPEPMANIPILLY 180  
 Db 121 IKGEMSRPGEVCSALNLCESLQKHLAELNHOKLESNKIPBLDMEVVAPEPMANIPILLY 180

QY 121 IKGEMSRPGEVCSALNLCESLQKHLAELNHOKLESNKIPBLDMEVVAPEPMANIPILLY 180  
 Db 121 IKGEMSRPGEVCSALNLCESLQKHLAELNHOKLESNKIPBLDMEVVAPEPMANIPILLY 180

QY 181 POGPRSKPQPKNDGVCODCIQWVTDIQTAVRTNSTFVQALVHVKEECDRLGPGMADI 240  
 Db 181 POGPRSKPQPKNDGVCODCIQWVTDIQTAVRTNSTFVQALVHVKEECDRLGPGMADI 240

QY 241 CKNYISQYSEIAIOMMMHMOQPKELCALVGFCDVEKEMPMOTLVPAKASKNVPALE 297  
 Db 241 CKNYISQYSEIAIOMMMHMOQPKELCALVGFCDVEKEMPMOTLVPAKASKNVPALE 297

QY 298 LVSEPIKHEVPASDVYCEVCEFLVKEVTKLIDNNKTEKELIDAFPMKSKLPKSLSEEC 357  
 Db 301 LVSEPIKHEVPASDVYCEVCEFLVKEVTKLIDNNKTEKELIDAFPMKSKLPKSLSEEC 357

QY 358 QEVVDYGSILSLILEEVSPELVCSMLHLCSTRLPALTVHTOPKDGFCVCKKLVG 417  
 Db 361 QEVVDYGSILSLILEEVSPELVCSMLHLCSTRLPALTVHTOPKDGFCVCKKLVG 417

QY 418 YIDRNLKSTKQELIAALEKGSFLPDPYQKQDFVAEYEPVLEILVEWMDPSFVCL 477  
 Db 421 YIDRNLKSTKQELIAALEKGSFLPDPYQKQDFVAEYEPVLEILVEWMDPSFVCL 477

QY 478 KIGACPSAHPPLGTREKCIWPSYWCQNTETAQCNVAEHCKRHAWN 524  
 Db 481 KIGACPSAHPPLGTREKCIWPSYWCQNTETAQCNVAEHCKRHAWN 524

RESULT 7  
 US-09-767-007A-2  
 ; Sequence 2, Application US/09767007A  
 ; Patent No. US20020072727SA1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: John S. O'Brien  
 ; APPLICANT: Yasuo Kishimoto  
 ; TITLE OF INVENTION: PHARMACEUTICAL COMPOSITIONS COMPRISING  
 ; FILE REFERENCE: MYELOS 2DC1C1  
 ; CURRENT APPLICATION NUMBER: US/09/767,007A  
 ; PRIOR FILING DATE: 2001-01-22  
 ; PRIOR APPLICATION NUMBER: 08/958,970  
 ; NUMBER OF SEQ ID NOS: 11  
 ; SOFTWARE: FastSeq for Windows Version 4.0  
 ; SEQ ID NO 2  
 ; LENGTH: 523  
 ; TYPE: PRT  
 ; ORGANISM: homo sapiens  
 ; US-09-767-007A-2

Query Match 98.9%; Score 2757.5; DB 9; Length 523;  
 Best Local Similarity 99.2%; Pred. No. 2.3e-236;  
 Matches 520; Conservative 2; Mismatches 1; Indels 1; Gaps 1;

QY 1 MYALFLASLILGALAGVTLGLKECTRGSAAVWCQNVKTASDCGAVKHCLQTVMNKPITYKS 60  
 Db 1 MYALFLASLILGALAGVTLGLKECTRGSAAVWCQNVKTASDCGAVKHCLQTVMNKPITYKS 60

QY 61 LPDCICKDVVTAAGMDLKDNAATEEELIVLEKTCMDLPRKPMNSASCKEIVDSYLPVILDI 120  
 Db 61 LPDCICKDVVTAAGMDLKDNAATEEELIVLEKTCMDLPRKPMNSASCKEIVDSYLPVILDI 120

QY 121 IKGEMSRPGEVCSALNLCESLQKHLAELNHOKLESNKIPBLDMEVVAPEPMANIPILLY 180  
 Db 121 IKGEMSRPGEVCSALNLCESLQKHLAELNHOKLESNKIPBLDMEVVAPEPMANIPILLY 180

QY 181 POGPRSKPQPKNDGVCODCIQWVTDIQTAVRTNSTFVQALVHVKEECDRLGPGMADI 240  
 Db 181 POGPRSKPQPKNDGVCODCIQWVTDIQTAVRTNSTFVQALVHVKEECDRLGPGMADI 240

QY 241 CKNYISQYSEIAIOMMMHMOQPKELCALVGFCDVEKEMPMOTLVPAKASKNVPALE 297  
 Db 240 CKNYISQYSEIAIOMMMHMOQPKELCALVGFCDVEKEMPMOTLVPAKASKNVPALE 297

QY 301 PIKHEVPASDVYCEVCEFLVKEVTKLIDNNKTEKELIDAFPMKSKLPKSLSEECOE 360  
 Db 300 PIKHEVPASDVYCEVCEFLVKEVTKLIDNNKTEKELIDAFPMKSKLPKSLSEECOE 359

QY 361 VDTYGSILSLILEEVSPELVCSMLHLCSTRLPALTVHTOPKDGFCVCKKLVG 417  
 Db 360 VDTYGSILSLILEEVSPELVCSMLHLCSTRLPALTVHTOPKDGFCVCKKLVG 417

QY 421 RNLKSTKQELIAALEKGSFLPDPYQKQDFVAEYEPVLEILVEWMDPSFVCL 480  
 Db 420 RNLKSTKQELIAALEKGSFLPDPYQKQDFVAEYEPVLEILVEWMDPSFVCL 479

QY 481 ACPSAHPPLGTREKCIWPSYWCQNTETAQCNVAEHCKRHAWN 524  
 Db 480 ACPSAHPPLGTREKCIWPSYWCQNTETAQCNVAEHCKRHAWN 523

RESULT 8  
 US-09-978-418-40  
 ; Sequence 40, Application US/09978418  
 ; Publication No. US20030118997A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Benjamin, Stephan  
 ; APPLICANT: Tanaka, Hiroaki  
 ; TITLE OF INVENTION: HUMAN CDNAS AND PROTEINS AND USES THEREOF  
 ; FILE REFERENCE: 142 US REG  
 ; CURRENT APPLICATION NUMBER: US/09/978,418  
 ; CURRENT FILING DATE: 2002-10-15  
 ; PRIOR APPLICATION NUMBER: 60/311,305  
 ; PRIOR FILING DATE: 2001-08-10  
 ; PRIOR APPLICATION NUMBER: 60/314,734  
 ; PRIOR FILING DATE: 2001-08-24  
 ; PRIOR APPLICATION NUMBER: 60/318,204  
 ; PRIOR FILING DATE: 2001-09-07  
 ; PRIOR APPLICATION NUMBER: 60/326,470  
 ; PRIOR FILING DATE: 2001-10-01  
 ; NUMBER OF SEQ ID NOS: 52  
 ; SOFTWARE: JPatent  
 ; SEQ ID NO 40  
 ; LENGTH: 479  
 ; TYPE: PRT  
 ; ORGANISM: Homo sapiens  
 ; US-09-978-418-40

Query Match 87.8%; Score 2449.5; DB 10; Length 479;  
 Best Local Similarity 99.4%; Pred. No. 5.1e-209;  
 Matches 469; Conservative 0; Mismatches 0; Indels 3; Gaps 1;

QY 1 MYALFLASLILGALAGVTLGLKECTRGSAAVWCQNVKTASDCGAVKHCLQTVMNKPITYKS 60  
 Db 1 MYALFLASLILGALAGVTLGLKECTRGSAAVWCQNVKTASDCGAVKHCLQTVMNKPITYKS 60

QY 61 LPDCICKDVVTAAGMDLKDNAATEEELIVLEKTCMDLPRKPMNSASCKEIVDSYLPVILDI 120  
 Db 61 LPDCICKDVVTAAGMDLKDNAATEEELIVLEKTCMDLPRKPMNSASCKEIVDSYLPVILDI 120

QY 121 IKGEMSRPGEVCSALNLCESLQKHLAELNHOKLESNKIPBLDMEVVAPEPMANIPILLY 180  
 Db 121 IKGEMSRPGEVCSALNLCESLQKHLAELNHOKLESNKIPBLDMEVVAPEPMANIPILLY 180

QY 181 POGPRSKPQPKNDGVCODCIQWVTDIQTAVRTNSTFVQALVHVKEECDRLGPGMADI 240

```
Db 181 PQDGRSRKQPKNDNDVQDCIQWTTDITQTAARTNSTFVQALVHVKESCRSLGPMADI 240
Qy 241 CKNYISQYSEIAIQMMHMM--OPKEICALVGFCEVYKEMPMQTLVPAKVASKNIPALYE 297
Db 241 CKNYISQYSEIAIQMMHMMQDOQPKKEICALVGFCEVYKEMPMQTLVPAKVASKNIPALYE 300
Qy 248 LVEPIKKKEVPAKSVYCEVCEFLVKEVTKLIDNNKTEKEIIDAPDKKCSKLPKSLSEEC 357
Db 301 LVEPIKKKEVPAKSVYCEVCEFLVKEVTKLIDNNKTEKEIIDAPDKKCSKLPKSLSEEC 360
Qy 358 QEVVDYSSILSILLESVPELVCSMLHLCSGTRLPALTVAHTOPKDGFCFCEVCKLVG 417
Db 361 QEVVDYSSILSILLESVPELVCSMLHLCSGTRLPALTVAHTOPKDGFCFCEVCKLVG 420
Qy 418 YLDRNLEKNSYKQETLALAEKGCSTLPDPYQKQCPQFAVEYPLIEILVEY 469
Db 421 YLDRNLEKNSYKQETLALAEKGCSTLPDPYQKQCPQFAVEYPLIEILVEY 472
```

RESULT 9  
US-10-205-194-176  
Sequence 176, Application US/10205194  
Publication No. US20030134301A1  
GENERAL INFORMATION:  
APPLICANT: Warner-Lambert Company  
APPLICANT: Lee, Kevin  
APPLICANT: Dixon, Alistair  
APPLICANT: Brooksbank, Robert  
APPLICANT: Pinnoch, Robert  
TITLE OF INVENTION: Identification and Use of Molecules Implicated in Pain  
FILE REFERENCE: WL-A-018201  
CURRENT APPLICATION NUMBER: US/10/205,194  
CURRENT FILING DATE: 5200-07-24  
PRIOR APPLICATION NUMBER: GB 0118354.0  
PRIOR FILING DATE: 2001-07-27  
NUMBER OF SEQ ID NOS: 177  
SOFTWARE: PatentIn Ver. 2.1  
SEQ ID NO: 176  
LENGTH: 554  
TYPE: PRT  
ORGANISM: Rattus sp.  
FEATURE:  
OTHER INFORMATION: Prosaposin  
US-10-205-194-176

```
Query Match 70.6%; Score 1969; DB 14; Length 554;  
Best Local Similarity 65.8%; Pred. No. 3,6e-166;  
Matches 365; Conservative 77; Mismatches 81; Indels 32; Gaps 3;  
Qy 1 MYALFLASLGLAGPLGLKECTRGSAVWCQVNTASDCAVYHCLQTVWNRPTYKS 60  
Db 1 MYALFLASLGLAGPLGLKECTRGSAVWCQVNTASDCAVYHCLQTVWNRPTYKS 60  
Qy 61 LPDCDCKOVVTAAGMLKDNATEEELVYLEKTCMPLPKNMSACKEIVDSYLVYIID 120  
Db 61 LPDCDCKOVVTAAGMLKDNATEEELVYLEKTCMPLPKNMSACKEIVDSYLVYIID 120  
Qy 121 IKGEMSRPEVCSALNLCESLOKHLAELNHQKLESNKIPELDMEVVAPEFMANIPLLY 180  
Db 121 IKGEMSRPEVCSALNLCESLOKHLAELNHQKLESNKIPELDMEVVAPEFMANIPLLY 180  
Qy 181 PODGRSRKQPKNDNDVQDCIQWTTDITQTAARTNSTFVQALVHVKESCRSLGPMADI 240  
Db 181 PODGRSRKQPKNDNDVQDCIQWTTDITQTAARTNSTFVQALVHVKESCRSLGPMADI 240  
Qy 241 CKNYISQYSEIAIQMMHMMQPKKEICALVGFCEVYKEMPMQTLVPAKVASKNIPALYE 297  
Db 241 CKNYISQYSEIAIQMMHMMQPKKEICALVGFCEVYKEMPMQTLVPAKVASKNIPALYE 297  
Qy 301 PIKKEHVPKSDVYCEVCEFLVKEVTKLIDNNKTEKEIIDAPDKKCSKLPKSLSEEC 360  
Db 301 PIKKEHVPKSDVYCEVCEFLVKEVTKLIDNNKTEKEIIDAPDKKCSKLPKSLSEEC 360  
Qy 361 QEVVDYSSILSILLESVPELVCSMLHLCSGTRLPALTVAHTOPKDGFCFCEVCKLVG 420  
Db 361 QEVVDYSSILSILLESVPELVCSMLHLCSGTRLPALTVAHTOPKDGFCFCEVCKLVG 420  
Qy 418 YLDRNLEKNSYKQETLALAEKGCSTLPDPYQKQCPQFAVEYPLIEILVEY 469  
Db 418 YLDRNLEKNSYKQETLALAEKGCSTLPDPYQKQCPQFAVEYPLIEILVEY 469  
Qy 421 YLDRNLEKNSYKQETLALAEKGCSTLPDPYQKQCPQFAVEYPLIEILVEY 472  
Db 421 YLDRNLEKNSYKQETLALAEKGCSTLPDPYQKQCPQFAVEYPLIEILVEY 472
```

```
Qy 361 VDTYSSILSILLESVPELVCSMLHLCSC-----TRLP----- 394  
Db 360 LVTFGRSLDVLMEHVENNFICGVISLCSANPNLVGTEDPAALIVSLPKEPPAPKPE 419  
Qy 395 -----ALTVHTOPKDGFCFCEVCKLVGYLDRNLEKNSYKQETLALAEKGCSTLPDPYQK 449  
Db 420 EPKQSALRAHP/PQKNGSFCEVCKLVYIEHNLKNSYKQETLALAEKGCSTLPDPYQK 479  
Qy 450 QCDQFAVEYPLIEILVEVNDPSFVCLKICACPSAKPPLIGTEKCTWGSYWCQNTETA 509  
Db 480 QCDQFAVEYPLIEILVEVNDPSFVCLKICACPSAKPPLIGTEKCTWGSYWCQNTETA 539  
Qy 510 AOCNAVEHCKRHWVN 524  
Db 540 ARCNADVCKRHWVN 554
```

RESULT 10  
US-10-276-162-1  
Sequence 1, Application US/10276162  
Publication No. US20030215822A1  
GENERAL INFORMATION:  
APPLICANT: GRIFFIN, Jennifer A.  
APPLICANT: YAO, Monique G.  
APPLICANT: BRUNS, Christopher M.  
APPLICANT: YUE, Henry  
APPLICANT: DELBEANE, Angelo M.  
APPLICANT: HAFALIA, April  
APPLICANT: PATTERSON, Chandra  
APPLICANT: POLICKY, Jennifer L.  
APPLICANT: TRIBOULEY, Catherine M.  
APPLICANT: BAUGHN, Mariah R.  
APPLICANT: NGUYEN, Danielle B.  
APPLICANT: LAL, Preeti  
APPLICANT: TANG, Y. Tom  
APPLICANT: JACKSON, Jennifer L.  
APPLICANT: LU, Dyrung Alma M.  
APPLICANT: BATRA, Saijeev  
APPLICANT: AU-YOUNG, Janice  
APPLICANT: REDDY, Roopa  
APPLICANT: AZIMZAI, Yalda  
TITLE OF INVENTION: SECRETED PROTEINS  
FILE REFERENCE: PI-0071 USN  
CURRENT APPLICATION NUMBER: US/10/276,162  
CURRENT FILING DATE: 2002-10-15  
PRIOR APPLICATION NUMBER: US01/11861  
PRIOR FILING DATE: 2001-04-11  
PRIOR APPLICATION NUMBER: 60/197,854  
PRIOR FILING DATE: 2000-04-14  
PRIOR APPLICATION NUMBER: 60/202,373  
PRIOR FILING DATE: 2000-05-04  
PRIOR APPLICATION NUMBER: 60/205,899  
PRIOR FILING DATE: 2000-05-18  
PRIOR APPLICATION NUMBER: 60/210,155  
PRIOR FILING DATE: 2000-06-01  
PRIOR APPLICATION NUMBER: 60/209,401  
NUMBER OF SEQ ID NOS: 28  
SOFTWARE: PERL Program  
SEQ ID NO: 1  
LENGTH: 521  
TYPE: PRT  
ORGANISM: Homo sapiens  
FEATURE:  
NAME/KEY: misc feature  
OTHER INFORMATION: Incyte ID No. US20030215822A1 7473577CD1  
US-10-276-162-1

Query Match 43.1%; Score 1201; DB 15; Length 521;  
Best Local Similarity 44.3%; Pred. No. 7.1e-98;  
Matches 237; Conservative 98; Mismatches 158; Indels 42; Gaps 10;

```

QY 3 ALFLASLLGALAGPVLTGKECTRGSAVWCQVKTASDCAVKGCLQTYWKNKPTYSKLP 62
DB 4 ALLILPBLGATRASPTSGPOEAKGSGTWQCDLQTARCAVGYCOGAVMKNPTAKSLP 63
QY 63 CDICKDVNTAAGDMLKDNATEEBELVYLEKTCMDLPKPNMSASCKEIVDSYLPVILDIK 122
DB 64 CDVCDIAAAGANGLNPDATESDILALVMKTCMLPSQESSAGCKMNVDAHSAILSMIR 123
QY 123 GEM-SRPGVCSALNLCESLOKHLAEINHQKLESNKIPELDMTEVAVPMANIPLLLYP 181
DB 124 GAPDSAPAOVCTALSLCEPLQRLHATL-----RLSKEDTFEAVAFPMANGPLTFHP 175
QY 182 QDPRSKQPKDNDGVCDQDCIQWVTDIQTAVRTNSFVQALVEHVKEECDRLGPMADIC 241
DB 176 RQAP-----EGALQDCQVROVSRLOEAVRSNLTLDL---NIQOCESLGFGLAVLC 224
QY 242 KNTISQYSEIALIOMMMHOPKEICALVGFCEVKEKEMQTLVPAK---VASKNVIPALEL 298
DB 225 KNVLFQFPVPADQALRLPPOELCRKGFCEELG-----APARLTQVAMDGVSLEL 277
QY 299 VEPKKEHPAKSDVYCEVCEFLVKEVTKLIDNNKTEKEILDAFDKMSKLPKSLSECO 358
DB 278 GLPRKQSEMOKAGVCEVCANVQKLDHLMNSSELMTTHALERVCSVMPASITKECI 337
QY 359 EVDVTYSSILILLEVSPELVCSMLHLCSGTRLPLATVH-----VTQPK-----DGGF 408
DB 338 ILVDITSPSLVQ-LVAKITPEKVKCFIRLC-GNRRRAVAVDAVAIVSPEDMAENQSGF 395
QY 409 CEVCKKLVGYLDNRNLEKNSTKOEILALEKSGSFLDPYQKCCDQFVAEYEPVILILE 468
DB 396 CNGCKRLITVSSHNSKSTKRDILVAFKGGCSILPLPYMIQCKHFVQYEBVILIESLKD 455
QY 469 VMDPSFVCLIKIGACPSAHKPLLTGTEKCIWGPSYWCQNTETAQCAVNEHCKRHVW 523
DB 456 WMDPVAACKVGAACHGPRTPPLIGTDQCALGSPFCWRSQEAALCNVAHGCQKHVW 510

RESULT 11
US-09-833-245-903
/ Sequence 903, Application US/09833245
/ Publication No. US20040010134A1
/ GENERAL INFORMATION:
/ APPLICANT: Human Genome Sciences, Inc.
/ TITLE OF INVENTION: Albumin Fusion Proteins
/ FILE REFERENCE: PFS46PCT
/ CURRENT APPLICATION NUMBER: US/09/833,245
/ PRIOR FILING DATE: 2001-04-12
/ PRIOR APPLICATION NUMBER: 60/229, 358
/ PRIOR FILING DATE: 2000-04-12
/ PRIOR APPLICATION NUMBER: 60/256, 931
/ PRIOR FILING DATE: 2000-12-21
/ PRIOR APPLICATION NUMBER: 60/199, 384
/ PRIOR FILING DATE: 2000-04-25
/ NUMBER OF SEQ ID NOS: 2267
/ SOFTWARE: PatentIn Ver. 2.1
/ SEQ ID NO 903
/ LENGTH: 531
/ TYPE: PRT
/ ORGANISM: Homo sapiens
US-09-833-245-903

```

```

Query Match 40.9%; Score 1139.5; DB 11; Length 531;
Best Local Similarity 43.2%; Pred. No. 2.2e-92;
Matches 231; Conservative 99; Mismatches 162; Indels 43; Gaps 11;
QY 3 ALFLASLLGALAGPVLTGKECTRGSAVWCQVKTASDCAVKGCLQTYWKNKPTYSKLP 62
DB 4 ALLILPBLGATRASPTSGPOEAKGSGTWQCDLQTARCAVGYCOGAVMKNPTAKSLP 63
QY 63 CDICKDVNTAAGDMLKDNATEEBELVYLEKTCMDLPKPNMSASCKEIVDSYLPVILDIK 122
DB 64 CDVCDIAAAGANGLNPDATESDILALVMKTCMLPSQESSAGCKMNVDAHSAILSMIR 123

```

```

QY 123 GEM-SRPGVCSALNLCESLOKHLAEINHQKLESNKIPELDMTEVAVPMANIPLLLYP 181
DB 124 GAPDSAPAOVCTALSLCEPLQRLHATL-----RLSKEDTFEAVAFPMANGPLTFHP 175
QY 182 QDPRSKQPKDNDGVCDQDCIQWVTDIQTAVRTNSFVQALVEHVKEECDRLGPMADIC 241
DB 176 RQAP-----EGALQDCQVROVSRLOEAVRSNLTLDL---NIQOCESLGFGLAVLC 224
QY 242 KNTISQYSEIALIOMMMHOPKEICALVGFCEVKEKEMQTLVPAK---VASKNVIPALEL 298
DB 225 KNVLFQFPVPADQALRLPPOELCRKGFCEELG-----APARLTQVAMDGVSLEL 277
QY 299 VEPKKEHPAKSDVYCEVCEFLVKEVTKLIDNNKTEKEILDAFDKMSKLPKSLSECO 358
DB 278 GLPRKQSEMOKAGVCEVCANVQKLDHLMNSSELMTTHALERVCSVMPASITKECI 337
QY 359 EVDVTYSSILILLEVSPELVCSMLHLCSGTRLPLATVH-----VTQPK-----DGGF 408
DB 338 ILVDITSPSLVQ-LVAKITPEKVKCFIRLC-GNRRRAVAVDAVAIVSPEDMAENQSGF 395
QY 409 CEVCKKLVGYLDNRNLEKNSTKOEILALEKSGSFLDPYQKCCDQFVAEYEPVILILE 468
DB 396 CNGCKRLITVSSHNSKSTKRDILVAFKGGCSILPLPYMIQCKHFVQYEBVILIESLKD 455
QY 469 VMDPSFVCLIKIGACPSAHKPLLTGTEKCIWGPSYWCQNTETAQCAVNEHCKRHVW 523
DB 456 WMDPVAACKVGAACHGPRTPPLIGTDQCALGSPFCWRSQEAALCNVAHGCQKHVW 509

RESULT 12
US-10-043-487-340
/ Sequence 340, Application US/10043487
/ Publication No. US20030055220A1
/ GENERAL INFORMATION:
/ APPLICANT: HYBRIGENICS
/ TITLE OF INVENTION: Protein-protein interactions between Shigella flexneri polypeptide
/ FILE REFERENCE: B4778A
/ CURRENT APPLICATION NUMBER: US/10/043,487
/ PRIOR FILING DATE: 2002-04-30
/ PRIOR APPLICATION NUMBER: US 60/261,130
/ PRIOR FILING DATE: 2001-01-12
/ NUMBER OF SEQ ID NOS: 561
/ SOFTWARE: PatentIn version 3.1
/ SEQ ID NO 340
/ LENGTH: 209
/ TYPE: PRT
/ ORGANISM: Shigella flexneri
US-10-043-487-340

```

```

Query Match 39.2%; Score 1093; DB 14; Length 209;
Best Local Similarity 100.0%; Pred. No. 7.5e-89;
Matches 209; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 151 OKQLESNKIPELDMTEVAVPMANIPLLLYPODGRSKQPKDNDGVCDQDCIQWVTDIQT 210
DB 1 OKQLESNKIPELDMTEVAVPMANIPLLLYPODGRSKQPKDNDGVCDQDCIQWVTDIQT 60
QY 211 AVRTNSTFVQALVEHVKEECDRLGPMADICKNTISQYSEIALIOMMMHOPKEICALVGF 270
DB 61 AVRTNSTFVQALVEHVKEECDRLGPMADICKNTISQYSEIALIOMMMHOPKEICALVGF 120
QY 271 CDEVKEMQTLVPAKVASKNVIPALELVEPIKKEHVPAKSDVYCEVCEFLVKEVTKLID 330
DB 121 CDEVKEMQTLVPAKVASKNVIPALELVEPIKKEHVPAKSDVYCEVCEFLVKEVTKLID 180
QY 331 NNKTEKEILDAFDKMSKLPKSLSECOE 359
DB 181 NNKTEKEILDAFDKMSKLPKSLSECOE 209

```

RESULT 13



Db 590 LIDMLITDPFKQECICVQKLCPRK-----TTATAM 618

RESULT 15  
US-10-108-260A-4529  
; Sequence 4529, Application US/10108260A  
; Publication No. US20040005560A1  
; GENERAL INFORMATION:  
; APPLICANT: HELIX RESEARCH INSTITUTE  
; TITLE OF INVENTION: No. US20040005560A1e1 full length cDNA  
; FILE REFERENCE: H1-A0106  
; CURRENT APPLICATION NUMBER: US/10/108,260A  
; CURRENT FILING DATE: 2002-03-27  
; NUMBER OF SEQ ID NOS: 5458  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 4529  
; LENGTH: 241  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-108-260A-4529

Query Match 19.5%; Score 545; DB 15; Length 241;  
Best Local Similarity 44.6%; Pred. No. 5.3e-40;  
Matches 107; Conservative 45; Mismatches 76; Indels 12; Gaps 4;

QY 293 IPALIEVPIKKEVPAKSDVYCEVCEFLVKEVTKLIDNNKTEKEILDAPDKMSKLPKS 352  
Db 4 VPSLEIGLPRKQSEMOMKAGVTCVCMNVOKLDHMLNSSELMITTHALERVCSPMPAS 63  
QY 353 LSECOEVVDYTGSSILSLILEVSPPELVCSMLHLCSTRLPALTVH-----VTQPK--- 404  
Db 64 ITKECTIIVDTYSPSLVQ-LVAKITPEKVKCFIRLC-GNRRRAVADAVAIVPSPEMDA 121  
QY 405 --DGGFCEVCKKLVGYLDNLEKSTKQELLALAEKGSFLDPYQKQCDQFVAEYEPVL 462  
Db 122 ENQGSFCNGCKRLLIVSSHNLESKSTIKRDLVAFKGCSTLPLPYMIQCKHFTQYEPL 181  
QY 463 IEILVEVMDPSFVCLKIGACPSAHKPLLGTEKCIWGPSYWCNTETAQCNAAVEHCKRHV 522  
Db 182 IESLKMMDPVAVCKKVGACHGPRFTPLGTQCCALGPSFCRSQDEAAKLCNAVQHCKRHV 241

Search completed: May 5, 2004, 13:28:22  
Job time : 56.7341 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 5, 2004, 13:10:37 ; Search time 22.8908 Seconds  
(without alignments)  
1181.786 Million cell updates/sec

Title: US-09-743-684A-1

Perfect score: 2789

Sequence: 1 MYAFLPLASLIGALAGPVL.....NTETRAQCNAYEHCKRHWN 524

Scoring table:

BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 389414 seqs, 51625971 residues

Total number of hits satisfying chosen parameters: 389414

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Issued Patents AA:\*  
1: /cgn2\_6/ptodata/2/1aa/5A.COMB.pep:\*  
2: /cgn2\_6/ptodata/2/1aa/5B.COMB.pep:\*  
3: /cgn2\_6/ptodata/2/1aa/6A.COMB.pep:\*  
4: /cgn2\_6/ptodata/2/1aa/6B.COMB.pep:\*  
5: /cgn2\_6/ptodata/2/1aa/PCTUS.COMB.pep:\*  
6: /cgn2\_6/ptodata/2/1aa/backfile1.pep:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2789	100.0	524	4	US-09-352-548-1
2	2766.5	99.2	523	1	US-08-100-247-2
3	2766.5	99.2	523	4	US-08-756-031-2
4	2759.5	98.9	523	1	US-08-232-513A-3
5	2757.5	98.9	523	1	US-08-483-146A-2
6	2757.5	98.9	523	1	US-08-484-594A-2
7	2757.5	98.9	523	4	US-09-076-258A-2
8	431	15.5	81	4	US-09-352-548-2
9	418.5	15.0	80	2	US-08-584-671-15
10	418.5	15.0	80	3	US-09-027-376-15
11	415.5	14.9	80	3	US-09-094-192-15
12	412	14.8	80	1	US-08-100-247-3
13	412	14.8	80	1	US-08-483-146A-3
14	412	14.8	80	1	US-08-232-513A-4
15	412	14.8	80	1	US-08-484-594A-3
16	412	14.8	80	4	US-09-076-258A-3
17	412	14.8	80	4	US-08-756-031-3
18	350.5	12.6	381	2	US-09-193-877-2
19	327	11.7	61	2	US-08-584-671-13
20	327	11.7	61	3	US-09-027-376-13
21	327	11.7	61	3	US-09-094-192-13
22	327	11.7	69	3	US-09-268-070-2
23	322	11.5	79	2	US-08-584-671-16
24	322	11.5	79	3	US-09-027-376-16
25	322	11.5	79	3	US-09-094-192-16
26	321	11.5	79	2	US-08-584-671-14
27	321	11.5	79	3	US-09-027-376-14

28	321	11.5	79	3	US-09-094-192-14	Sequence 14, Appl
29	268.5	9.6	257	3	US-08-596-684F-7	Sequence 7, Appl
30	239.5	8.6	60	2	US-08-584-671-12	Sequence 12, Appl
31	239.5	8.6	60	3	US-09-027-376-12	Sequence 12, Appl
32	239.5	8.6	60	3	US-09-268-070-4	Sequence 4, Appl
33	239.5	8.6	60	3	US-09-094-192-12	Sequence 12, Appl
34	238.5	8.6	68	3	US-09-268-070-3	Sequence 3, Appl
35	208	7.5	67	3	US-09-268-070-1	Sequence 1, Appl
36	168.5	6.0	181	4	US-08-848-580-12	Sequence 12, Appl
37	168.5	6.0	181	4	US-08-488-123-12	Sequence 12, Appl
38	126.5	4.5	514	4	US-09-107-532A-6282	Sequence 6282, Ap
39	115.5	4.1	1002	4	US-09-762-724-4	Sequence 4, Appl
40	113	4.1	1027	4	US-09-762-724-8	Sequence 8, Appl
41	113	4.1	1027	4	US-08-973-462-8	Sequence 8, Appl
42	112.5	4.0	1014	4	US-09-762-724-2	Sequence 2, Appl
43	111	4.0	22	1	US-08-100-247-1	Sequence 1, Appl
44	111	4.0	22	1	US-08-483-146A-1	Sequence 1, Appl
45	111	4.0	22	1	US-08-483-146A-1	Sequence 11, Appl

ALIGNMENTS

RESULT 1									
US-09-352-548-1									
Sequence 1, Application US/09352548									
Patent No. 6500431									
GENERAL INFORMATION:									
APPLICANT: Gill, Parkash S.									
APPLICANT: Parkash S. Gill, M.D., Inc.									
TITLE OF INVENTION: No. 6500431el Inhibitors of Angiogenesis and Tumor Growth									
FILE REFERENCE: 017968-000410US									
CURRENT APPLICATION NUMBER: US/09/352,548									
CURRENT FILING DATE: 1999-07-12									
EARLIER APPLICATION NUMBER: US 60/092,647									
EARLIER FILING DATE: 1998-07-13									
NUMBER OF SEQ ID NOS: 59									
SOFTWARE: PatentIn Ver. 2.1									
SEQ ID NO 1									
LENGTH: 524									
TYPE: PRT									
ORGANISM: Homo sapiens									
FEATURE:									
OTHER INFORMATION: prosaposin									
FEATURE:									
NAME/KEY: PEPTIDE									
LOCATION: (195)..(275)									
OTHER INFORMATION: Saposin B									
US-09-352-548-1									
Query Match									
Best Local Similarity 100.0%; Pred. No. 1.3e-242;									
Matches 524; Conservative 0; Mismatches 0; Indels 0; Gaps 0;									
QY	1	MYAFLPLASLIGALAGPVLGLKCTGSAVQCNVTAPOGAVKCLQTVNKKPTVKS	60						
DB	1	MYAFLPLASLIGALAGPVLGLKCTGSAVQCNVTAPOGAVKCLQTVNKKPTVKS	60						
QY	61	LPCDICKDVVTAADMDLKDNTAEELIIVLEKTDMDLKPKNMSASCKEIVDSYLPVLDI	120						
DB	61	LPCDICKDVVTAADMDLKDNTAEELIIVLEKTDMDLKPKNMSASCKEIVDSYLPVLDI	120						
QY	121	IKGMSRGEVCSALNLCESIQKHLAEINHOKELENSKIPFLDTEVVAAPPMANIPILLY	180						
DB	121	IKGMSRGEVCSALNLCESIQKHLAEINHOKELENSKIPFLDTEVVAAPPMANIPILLY	180						
QY	181	PODGRSKPOPKNDGVCCDCTQAVTIDIQVAVTNSITFVQALVBEKBCDRLPGMADI	240						
DB	181	PODGRSKPOPKNDGVCCDCTQAVTIDIQVAVTNSITFVQALVBEKBCDRLPGMADI	240						
QY	241	CKNYSISYSEIAIQMMHMQPKIECALVGFCEVKEPMQTLVPAKVASKNVIPAELVE	300						
DB	241	CKNYSISYSEIAIQMMHMQPKIECALVGFCEVKEPMQTLVPAKVASKNVIPAELVE	300						

QY 301 PIKHEVPAKSDVYCEVCEFLVKEVTKLIDNNKTEKEILDAFPMKSKLPKSLSECCQEV 360  
Db 301 PIKHEVPAKSDVYCEVCEFLVKEVTKLIDNNKTEKEILDAFPMKSKLPKSLSECCQEV 360  
QY 361 VDTYGSIIISILLESVPELVCSMLHLCSTRLPALTVHTQPKDGFCEVCCKLVGYLD 420  
Db 361 VDTYGSIIISILLESVPELVCSMLHLCSTRLPALTVHTQPKDGFCEVCCKLVGYLD 420  
QY 421 RNLEKSTKOEIIAALKEGCSFLPDYQKCCDFVAEYEPVLEIIVEMDPSFVCLKIG 480  
Db 421 RNLEKSTKOEIIAALKEGCSFLPDYQKCCDFVAEYEPVLEIIVEMDPSFVCLKIG 480  
QY 481 ACPSAHKPLGTGTEKCIWGPSYWCNTETAACNAVHCHCRHWYN 524  
Db 481 ACPSAHKPLGTGTEKCIWGPSYWCNTETAACNAVHCHCRHWYN 524

RESULT 2  
US-08-100-247-2  
Sequence 2, Application US/08100247  
Patent No. 5571787  
GENERAL INFORMATION:  
APPLICANT: O'BRIEN, JOHN S.  
APPLICANT: KISHIMOTO, YASUO  
TITLE OF INVENTION: PROSAPOSIN AS A NEUTROTROPIC FACTOR  
NUMBER OF SEQUENCES: 5  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: KNOBBE, MARTENS, OLSON AND BEAR  
STREET: 620 NEWPORT CENTER DRIVE SIXTEENTH FLOOR  
CITY: NEWPORT BEACH  
STATE: CA  
COUNTRY: USA  
ZIP: 92660  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent In Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/100,247  
FILING DATE: 19930730  
CLASSIFICATION: 514  
ATTORNEY/AGENT INFORMATION:  
NAME: Israelien, Ned A.  
REGISTRATION NUMBER: 29,655  
REFERENCE/DOCKET NUMBER: OBR1EN.002A  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 619-235-8550  
TELEFAX: 619-235-0176  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 523 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
MOLECULE TYPE: linear  
TOPOLOGY: linear  
HYDROTHERMAL: NO  
ANTI-SENSE: NO  
FRAGMENT TYPE: N-terminal  
IMMEDIATE SOURCE:  
CLONE: PROSAPOSIN  
US-08-100-247-2

Query Match 99.2%; Score 2766.5; DB 1; Length 523;  
Best Local Similarity 99.4%; Pred. No. 1.4e-240;  
Matches 521; Conservative 2; Mismatches 0; Indels 1; Gaps 1;

QY 1 MYAFLLASLLGALAGPVGLKCTGSAVWCNVTTADGAVHGLCTVNNKPTVVS 60  
Db 1 MYAFLLASLLGALAGPVGLKCTGSAVWCNVTTADGAVHGLCTVNNKPTVVS 60  
QY 61 LPCDICKDVVTAAGMDLKNDAATEEELIVYEKTCMDMLPKRNMSASCKEIVDSYLPVILDI 120

Db 61 LPCDICKDVVTAAGMDLKNDAATEEELIVYEKTCMDMLPKRNMSASCKEIVDSYLPVILDI 120  
QY 121 IKGEMSRGCEVCSALNICESEIQRKLAFLNOKQLESNKIEBELDTEVVAFFMANIPALLY 180  
Db 121 IKGEMSRGCEVCSALNICESEIQRKLAFLNOKQLESNKIEBELDTEVVAFFMANIPALLY 180  
QY 181 PDDGPRSKPQKNDGVQCDCIOWVTDIQTAVRTNSTFVOALVEHVEECDRIGPGMADI 240  
Db 181 PDDGPRSKPQKNDGVQCDCIOWVTDIQTAVRTNSTFVOALVEHVEECDRIGPGMADI 240  
QY 241 CKNTYSQYSEIATIMMMHMQPKICALVGFCEVKEKMPQTLVPAKASNVIPALELVE 300  
Db 241 CKNTYSQYSEIATIMMMHMQPKICALVGFCEVKEKMPQTLVPAKASNVIPALELVE 300  
QY 301 PIKHEVPAKSDVYCEVCEFLVKEVTKLIDNNKTEKEILDAFPMKSKLPKSLSECCQEV 360  
Db 301 PIKHEVPAKSDVYCEVCEFLVKEVTKLIDNNKTEKEILDAFPMKSKLPKSLSECCQEV 360  
QY 361 VDTYGSIIISILLESVPELVCSMLHLCSTRLPALTVHTQPKDGFCEVCCKLVGYLD 420  
Db 361 VDTYGSIIISILLESVPELVCSMLHLCSTRLPALTVHTQPKDGFCEVCCKLVGYLD 420  
QY 421 RNLEKSTKOEIIAALKEGCSFLPDYQKCCDFVAEYEPVLEIIVEMDPSFVCLKIG 480  
Db 421 RNLEKSTKOEIIAALKEGCSFLPDYQKCCDFVAEYEPVLEIIVEMDPSFVCLKIG 480  
QY 481 ACPSAHKPLGTGTEKCIWGPSYWCNTETAACNAVHCHCRHWYN 524  
Db 481 ACPSAHKPLGTGTEKCIWGPSYWCNTETAACNAVHCHCRHWYN 524

RESULT 3  
US-08-756-031-2  
Sequence 2, Application US/08756031  
Patent No. 6590074  
GENERAL INFORMATION:  
APPLICANT: O'BRIEN, JOHN S.  
APPLICANT: KISHIMOTO, YASUO  
TITLE OF INVENTION: PROSAPOSIN AS A NEUTROTROPIC FACTOR  
NUMBER OF SEQUENCES: 5  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: KNOBBE, MARTENS, OLSON AND BEAR  
STREET: 620 NEWPORT CENTER DRIVE SIXTEENTH FLOOR  
CITY: NEWPORT BEACH  
STATE: CA  
COUNTRY: USA  
ZIP: 92660  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent In Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/756,031  
FILING DATE: 26-NOV-1996  
CLASSIFICATION: 514  
ATTORNEY/AGENT INFORMATION:  
NAME: Israelien, Ned A.  
REGISTRATION NUMBER: 29,655  
REFERENCE/DOCKET NUMBER: OBR1EN.002A  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 619-235-8550  
TELEFAX: 619-235-0176  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 523 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear



MOLECULE TYPE: peptide  
HYPOTHETICAL: NO  
ANTI-SENSE: NO  
FRAGMENT TYPE: N-terminal  
IMMEDIATE SOURCE:  
CLONE: PROSAPOSIN  
US-08-756-031-2

Query Match 99.2%; Score 2766.5; DB 4; Length 523;  
Best Local Similarity 99.4%; Pred. No. 1.4e-240;  
Matches 521; Conservative 2; Mismatches 0; Indels 1; Gaps 1;

QY 1 MYALFLASLIGALAGVPLGLKECTRGSAVWCQNVKTASDCGAVGKCLQTVMNKPPTYS 60  
DB 1 MYALFLASLIGALAGVPLGLKECTRGSAVWCQNVKTASDCGAVGKCLQTVMNKPPTYS 60  
QY 61 LPDCICKDVTAAGDMLKDNATEEELVYLEKTCMDLPKPMNASCKEIVDSYLPVLDI 120  
DB 61 LPDCICKDVTAAGDMLKDNATEEELVYLEKTCMDLPKPMNASCKEIVDSYLPVLDI 120  
QY 121 IKGEMSRGGEVCSALNLCESLQKHLAEINHOQKLESNKIPELDMEVVAPEPMANIPLLLY 180  
DB 121 IKGEMSRGGEVCSALNLCESLQKHLAEINHOQKLESNKIPELDMEVVAPEPMANIPLLLY 180  
QY 181 PODGPRSRKPOKNDGVCODCIQWVTDIQTAVRTNSTFVQALVEHKECDRLGPGMADI 240  
DB 181 PODGPRSRKPOKNDGVCODCIQWVTDIQTAVRTNSTFVQALVEHKECDRLGPGMADI 240  
QY 241 CKNYISOYSEIAIOMMMHMQPKICALVGFCDVEKEMQTLVPAKVASKNVIPALETVE 300  
DB 241 CKNYISOYSEIAIOMMMHMQPKICALVGFCDVEKEMQTLVPAKVASKNVIPALETVE 300  
QY 301 PIKKEVPAKSDVYCEVCEFLVKEVTYKLIIDNNKTEKEILDAFDKMSCLKPSLSSECOEV 360  
DB 301 PIKKEVPAKSDVYCEVCEFLVKEVTYKLIIDNNKTEKEILDAFDKMSCLKPSLSSECOEV 360  
QY 361 VDTYSSILSLILEEVSPELVCSMLHLCGSTRLPALTVAHTVOPKDGGECEVCKLVGYLD 420  
DB 361 VDTYSSILSLILEEVSPELVCSMLHLCGSTRLPALTVAHTVOPKDGGECEVCKLVGYLD 420  
QY 421 RNLEKNSIKQEIILALEKGCFLPDPPYKQCDQFAAEVEPVLIELVEMDPSFVCLKIG 480  
DB 421 RNLEKNSIKQEIILALEKGCFLPDPPYKQCDQFAAEVEPVLIELVEMDPSFVCLKIG 480  
QY 481 ACPSAHKPDLGTEKCIWGPSTYWCNTETAAOCNAVEHCKRHVYN 524  
DB 481 ACPSAHKPDLGTEKCIWGPSTYWCNTETAAOCNAVEHCKRHVYN 524

RESULT 4  
US-08-232-513A-3  
Sequence 3, Application US/08232513A  
Patent No. 5700909  
GENERAL INFORMATION:  
APPLICANT: O'Brien, John S.  
TITLE OF INVENTION: Prosapoin and Cytokine-Derived Peptides  
TITLE OF INVENTION: as Therapeutic Agents  
NUMBER OF SEQUENCES: 20  
CORRESPONDENCE ADDRESSES:  
ADDRESSES: Campbell & Flores LLP  
STREET: 4370 La Jolla Village Drive, Suite 700  
CITY: San Diego  
STATE: California  
COUNTRY: USA  
ZIP: 92122  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/232,513A  
FILING DATE: 21-APR-1994

CLASSIFICATION: 514  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/100,247  
FILING DATE: 30-JUL-1993  
ATTORNEY/AGENT INFORMATION:  
NAME: Campbell, Cathryn A.  
REGISTRATION NUMBER: 31,815  
REFERENCE/DOCKET NUMBER: P-UD 1643  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (619) 535-9001  
TELEFAX: (619) 535-8949  
INFORMATION FOR SEQ ID NO: 3:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 523 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
FEATURE:  
NAME/KEY: Protein  
LOCATION: 1..523  
OTHER INFORMATION: /label= Hum\_prosapoin  
US-08-232-513A-3

Query Match 98.9%; Score 2759.5; DB 1; Length 523;  
Best Local Similarity 99.2%; Pred. No. 5.9e-240;  
Matches 520; Conservative 2; Mismatches 1; Indels 1; Gaps 1;

QY 1 MYALFLASLIGALAGVPLGLKECTRGSAVWCQNVKTASDCGAVGKCLQTVMNKPPTYS 60  
DB 1 MYALFLASLIGALAGVPLGLKECTRGSAVWCQNVKTASDCGAVGKCLQTVMNKPPTYS 60  
QY 61 LPDCICKDVTAAGDMLKDNATEEELVYLEKTCMDLPKPMNASCKEIVDSYLPVLDI 120  
DB 61 LPDCICKDVTAAGDMLKDNATEEELVYLEKTCMDLPKPMNASCKEIVDSYLPVLDI 120  
QY 121 IKGEMSRGGEVCSALNLCESLQKHLAEINHOQKLESNKIPELDMEVVAPEPMANIPLLLY 180  
DB 121 IKGEMSRGGEVCSALNLCESLQKHLAEINHOQKLESNKIPELDMEVVAPEPMANIPLLLY 180  
QY 181 PODGPRSRKPOKNDGVCODCIQWVTDIQTAVRTNSTFVQALVEHKECDRLGPGMADI 240  
DB 181 PODGPRSRKPOKNDGVCODCIQWVTDIQTAVRTNSTFVQALVEHKECDRLGPGMADI 240  
QY 241 CKNYISOYSEIAIOMMMHMQPKICALVGFCDVEKEMQTLVPAKVASKNVIPALETVE 300  
DB 241 CKNYISOYSEIAIOMMMHMQPKICALVGFCDVEKEMQTLVPAKVASKNVIPALETVE 300  
QY 301 PIKKEVPAKSDVYCEVCEFLVKEVTYKLIIDNNKTEKEILDAFDKMSCLKPSLSSECOEV 360  
DB 301 PIKKEVPAKSDVYCEVCEFLVKEVTYKLIIDNNKTEKEILDAFDKMSCLKPSLSSECOEV 360  
QY 361 VDTYSSILSLILEEVSPELVCSMLHLCGSTRLPALTVAHTVOPKDGGECEVCKLVGYLD 420  
DB 361 VDTYSSILSLILEEVSPELVCSMLHLCGSTRLPALTVAHTVOPKDGGECEVCKLVGYLD 420  
QY 421 RNLEKNSIKQEIILALEKGCFLPDPPYKQCDQFAAEVEPVLIELVEMDPSFVCLKIG 480  
DB 421 RNLEKNSIKQEIILALEKGCFLPDPPYKQCDQFAAEVEPVLIELVEMDPSFVCLKIG 480  
QY 481 ACPSAHKPDLGTEKCIWGPSTYWCNTETAAOCNAVEHCKRHVYN 524  
DB 481 ACPSAHKPDLGTEKCIWGPSTYWCNTETAAOCNAVEHCKRHVYN 524

RESULT 5  
US-08-483-146A-2  
Sequence 2, Application US/08483146A  
Patent No. 5696080  
GENERAL INFORMATION:  
APPLICANT: O'Brien, John S.  
TITLE OF INVENTION: PHARMACEUTICAL COMPOSITIONS  
TITLE OF INVENTION: COMPRISING PROSAPOSIN AND NEUROTROPHIC PEPTIDES DERIVED

```

: TITLE OF INVENTION: THEREFROM
: NUMBER OF SEQUENCES: 11
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Knobbe, Martens, Olson and Bear
: STREET: 620 Newport Center Blvd. 16th Floor
: CITY: Newport Beach
: STATE: CA
: COUNTRY: USA
: ZIP: 92660
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Diskette
: COMPUTER: IBM Compatible
: OPERATING SYSTEM: DOS
: SOFTWARE: FastSeq for Windows Version 2.0
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/483,146A
: FILING DATE: 07-JUN-1995
: CLASSIFICATION: 514
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER:
: FILING DATE:
: ATTORNEY/AGENT INFORMATION:
: NAME: Israelson, Ned A
: REGISTRATION NUMBER: 29,655
: REFERENCE/DOCKET NUMBER: MYELOS.002DV1
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: 619-235-8550
: TELEFAX: 619-235-0176
:
: TELEX:
: INFORMATION FOR SEQ ID NO: 2:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 523 amino acids
: TYPE: amino acid
: STRANDEDNESS: single
: TOPOLOGY: linear
: MOLECULE TYPE: peptide
: FRAGMENT TYPE: N-terminal
: US-08-483-146A-2
:
: Query Match      98.9%; Score 2757.5; DB 1; Length 523;
: Best Local Similarity 99.2%; Pred. No. 8.9e-240;
: Matches 520; Conservative 2; Mismatches 1; Indels 1; Gaps 1;
:
: QY 1 MYALFLASLGAALAGPVLGLKECTRGSAVWCQNVKTASDCGAVHCLQTVWVKPTYS 60
: DB 1 MYALFLASLGAALAGPVLGLKECTRGSAVWCQNVKTASDCGAVHCLQTVWVKPTYS 60
: QY 61 LPDICKDVVTAAGDMLKDNATEEELVYLEKTCDWLPRPNMSASCKEIVDSYLPVILDI 120
: DB 61 LPDICKDVVTAAGDMLKDNATEEELVYLEKTCDWLPRPNMSASCKEIVDSYLPVILDI 120
: QY 121 IKGEMSRPGEVCSALNLCESLQKHLAELNHQKQLESNKIPELDMTEVVAAPFMANIPILLY 180
: DB 121 IKGEMSRPGEVCSALNLCESLQKHLAELNHQKQLESNKIPELDMTEVVAAPFMANIPILLY 180
: QY 121 IKGEMSRPGEVCSALNLCESLQKHLAELNHQKQLESNKIPELDMTEVVAAPFMANIPILLY 180
: DB 121 IKGEMSRPGEVCSALNLCESLQKHLAELNHQKQLESNKIPELDMTEVVAAPFMANIPILLY 180
: QY 181 POGPRSKPOPKD-GDVCODCIOMVTDIOTAVRTNSTFVQALVHVHKECDRLGPGMADI 240
: DB 181 POGPRSKPOPKD-GDVCODCIOMVTDIOTAVRTNSTFVQALVHVHKECDRLGPGMADI 240
: QY 241 CKNYISQVSEIAIQMMHMQPEICAVGFCDEVKEMQTLVPKVASKNVIVPALIVE 300
: DB 241 CKNYISQVSEIAIQMMHMQPEICAVGFCDEVKEMQTLVPKVASKNVIVPALIVE 300
: QY 301 PIKKEHVAKSDVYCEVCEFLVKEVTKLIDNNKTEKEILDAFDKMSCLPKSLSEECOE 360
: DB 301 PIKKEHVAKSDVYCEVCEFLVKEVTKLIDNNKTEKEILDAFDKMSCLPKSLSEECOE 360
: QY 360 VDTGSSSLTSLLEEVSPELVCSMLHLCSGTRLPALTVHVTQPKDGGCEVCKKLVGLD 420
: DB 360 VDTGSSSLTSLLEEVSPELVCSMLHLCSGTRLPALTVHVTQPKDGGCEVCKKLVGLD 420
: QY 421 RNLEKNTKQETILALEKSGSFLDPYKOCODQVAVAEVPELITLVWMDPSVCLKIG 480
: DB 421 RNLEKNTKQETILALEKSGSFLDPYKOCODQVAVAEVPELITLVWMDPSVCLKIG 480

```

```

: DB 420 RNLEKNTKQETILALEKSGSFLDPYKOCODQVAVAEVPELITLVWMDPSVCLKIG 479
: QY 481 ACPSAHKRPLGTCEKICWPSYWCNTETAAQCNVAHECKSHWN 524
: DB 480 ACPSAHKRPLGTCEKICWPSYWCNTETAAQCNVAHECKSHWN 523
:
: RESULT 6
: US-08-484-594A-2
: Sequence 2, Application US/08484594A
: Patent No. 5714459
: GENERAL INFORMATION:
: APPLICANT: O'Brien, John S.
: APPLICANT: Kishimoto, Yasuo
: TITLE OF INVENTION: USE OF PROSAPOSIN AND NEUROTROPHIC PEPTIDES
: NUMBER OF SEQUENCES: 11
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Knobbe, Martens, Olson and Bear
: STREET: 620 Newport Center Drive, Sixteenth Floor
: CITY: Newport Beach
: STATE: CA
: COUNTRY: USA
: ZIP: 92660
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Diskette
: COMPUTER: IBM Compatible
: OPERATING SYSTEM: DOS
: SOFTWARE: FastSeq for Windows Version 2.0
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/484,594A
: FILING DATE: 07-JUN-1995
: CLASSIFICATION: 514
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: 08/100,247
: FILING DATE: 30-JUL-1993
: ATTORNEY/AGENT INFORMATION:
: NAME: Israelson, Ned A
: REGISTRATION NUMBER: 29,655
: REFERENCE/DOCKET NUMBER: MYELOS.002DV2
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: 619-235-8550
: TELEFAX: 619-235-0176
:
: TELEX:
: INFORMATION FOR SEQ ID NO: 2:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 523 amino acids
: TYPE: amino acid
: STRANDEDNESS: single
: TOPOLOGY: linear
: MOLECULE TYPE: peptide
: FRAGMENT TYPE: N-terminal
: US-08-484-594A-2
:
: Query Match      98.9%; Score 2757.5; DB 1; Length 523;
: Best Local Similarity 99.2%; Pred. No. 8.9e-240;
: Matches 520; Conservative 2; Mismatches 1; Indels 1; Gaps 1;
:
: QY 1 MYALFLASLGAALAGPVLGLKECTRGSAVWCQNVKTASDCGAVHCLQTVWVKPTYS 60
: DB 1 MYALFLASLGAALAGPVLGLKECTRGSAVWCQNVKTASDCGAVHCLQTVWVKPTYS 60
: QY 61 LPDICKDVVTAAGDMLKDNATEEELVYLEKTCDWLPRPNMSASCKEIVDSYLPVILDI 120
: DB 61 LPDICKDVVTAAGDMLKDNATEEELVYLEKTCDWLPRPNMSASCKEIVDSYLPVILDI 120
: QY 121 IKGEMSRPGEVCSALNLCESLQKHLAELNHQKQLESNKIPELDMTEVVAAPFMANIPILLY 180
: DB 121 IKGEMSRPGEVCSALNLCESLQKHLAELNHQKQLESNKIPELDMTEVVAAPFMANIPILLY 180
: QY 181 POGPRSKPOPKD-GDVCODCIOMVTDIOTAVRTNSTFVQALVHVHKECDRLGPGMADI 240
: DB 181 POGPRSKPOPKD-GDVCODCIOMVTDIOTAVRTNSTFVQALVHVHKECDRLGPGMADI 239

```

QY 241 CKNYISQYSEIAIQMMHMPKEICLVGFCDEVKEMPMQTLVPAKVASKNVIPLALVE 300  
DB 240 CKNYISQYSEIAIQMMHMPKEICLVGFCDEVKEMPMQTLVPAKVASKNVIPLALVD 299  
QY 301 PIKKEVPAKSDVYCEVCEFLVKEVTYKLIIDNNKTEKEILDAFDKXCSKLPKSLSECCOV 360  
DB 300 PIKKEVPAKSDVYCEVCEFLVKEVTYKLIIDNNKTEKEILDAFDKXCSKLPKSLSECCOV 359  
QY 361 VDTYSSILSLILEEVSPELVCSMLHLCSTRLPALTYHTVTPKDGFCCEVCCKLVGLD 420  
DB 360 VDTYSSILSLILEEVSPELVCSMLHLCSTRLPALTYHTVTPKDGFCCEVCCKLVGLD 419  
QY 421 RNLEKNSKQEBILALEKGCSEFLPDYKQCDQFVAEYEPVLIETLVEMDPSFVCLKIG 480  
DB 420 RNLEKNSKQEBILALEKGCSEFLPDYKQCDQFVAEYEPVLIETLVEMDPSFVCLKIG 479  
QY 481 ACPSAHKPLGTCKICWGPSYWCNTETPAQCNAVEHCKRHVN 524  
DB 480 ACPSAHKPLGTCKICWGPSYWCNTETPAQCNAVEHCKRHVN 523

RESULT 7  
US-09-076-258A-2

; Sequence 2, Application US/09076258A  
; Patent No. 6559124

## GENERAL INFORMATION:

APPLICANT: O'Brien, John S.

APPLICANT: Kishimoto, Yasuo

TITLE OF INVENTION: PHARMACEUTICAL COMPOSITIONS

TITLE OF INVENTION: COMPRISING PROSAPOSIN AND NEUTROPHIC PEPTIDES DERIVED

NUMBER OF SEQUENCES: 11

CORRESPONDENCE ADDRESS:

ADDRESSEE: Knobb, Martens, Olson and Bear

STREET: 620 Newport Center Blvd. 16th Floor

CITY: Newport Beach

STATE: CA

COUNTRY: USA

ZIP: 92660

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette

COMPUTER: IBM Compatible

OPERATING SYSTEM: DOS

SOFTWARE: FastSeq for Windows Version 2.0

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/076,258A

FILING DATE:

CLASSIFICATION:

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 08/958,970

FILING DATE: 28-OCT-97

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 08/483,146

FILING DATE: 07-JUN-1995

APPLICATION NUMBER: 08/100,247

FILING DATE: 30-JUL-1993

ATTORNEY/AGENT INFORMATION:

NAME: Israel, Ned A

REGISTRATION NUMBER: 29,655

REFERENCE/DOCKET NUMBER: MYELOS.2DV1C2

TELEPHONE: 619-235-8550

TELEFAX: 619-235-0176

TELEX:

INFORMATION FOR SEQ ID NO: 2:

SEQUENCE CHARACTERISTICS:

LENGTH: 523 amino acids

TYPE: amino acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: peptide

FRAGMENT TYPE: N-terminal

## US-09-076-258A-2

Query Match 98.9%; Score 2757.5; DB 4; Length 523;

Best Local Similarity 99.2%; Pred. No. 8.9e-240;

Matches 520; Conservative 2; Mismatches 1; Indels 1; Gaps 1;

QY 1 MYALFLASLILGAALAGVPLGLKECTRGSAVWCQNVKPAASDCGAVKCLQTVNKPVTYS 60  
DB 1 MYALFLASLILGAALAGVPLGLKECTRGSAVWCQNVKPAASDCGAVKCLQTVNKPVTYS 60  
QY 61 LPDCIKDVVTAAGDMLKDNATEEELIYLEKTCMDLPKPMNASCKEIVDSYLPVILI 120  
DB 61 LPDCIKDVVTAAGDMLKDNATEEELIYLEKTCMDLPKPMNASCKEIVDSYLPVILI 120  
QY 121 IKGEMSRGCVCSALNLCESIQKHAELNHNKQLESNKIPBLDMTEVVAAPPMANIPLLY 180  
DB 121 IKGEMSRGCVCSALNLCESIQKHAELNHNKQLESNKIPBLDMTEVVAAPPMANIPLLY 180  
QY 181 PDGPRSRKPRKNDVDCODCIQWVTDIQTAVRTNSTFVQALVEHVEKCDRLGPMADI 240  
DB 181 PDGPRSRKPRKNDVDCODCIQWVTDIQTAVRTNSTFVQALVEHVEKCDRLGPMADI 239  
QY 241 CKNYISQYSEIAIQMMHMPKEICLVGFCDEVKEMPMQTLVPAKVASKNVIPLALVE 300  
DB 240 CKNYISQYSEIAIQMMHMPKEICLVGFCDEVKEMPMQTLVPAKVASKNVIPLALVD 299  
QY 301 PIKKEVPAKSDVYCEVCEFLVKEVTYKLIIDNNKTEKEILDAFDKXCSKLPKSLSECCOV 360  
DB 300 PIKKEVPAKSDVYCEVCEFLVKEVTYKLIIDNNKTEKEILDAFDKXCSKLPKSLSECCOV 359  
QY 361 VDTYSSILSLILEEVSPELVCSMLHLCSTRLPALTYHTVTPKDGFCCEVCCKLVGLD 420  
DB 360 VDTYSSILSLILEEVSPELVCSMLHLCSTRLPALTYHTVTPKDGFCCEVCCKLVGLD 419  
QY 421 RNLEKNSKQEBILALEKGCSEFLPDYKQCDQFVAEYEPVLIETLVEMDPSFVCLKIG 480  
DB 420 RNLEKNSKQEBILALEKGCSEFLPDYKQCDQFVAEYEPVLIETLVEMDPSFVCLKIG 479  
QY 481 ACPSAHKPLGTCKICWGPSYWCNTETPAQCNAVEHCKRHVN 524  
DB 480 ACPSAHKPLGTCKICWGPSYWCNTETPAQCNAVEHCKRHVN 523

## RESULT 8

US-09-352-548-2

; Sequence 2, Application US/09352548

; Patent No. 6500431

GENERAL INFORMATION:

APPLICANT: Gill, Parkash S.

APPLICANT: Parkash S. Gill, M.D., Inc.

TITLE OF INVENTION: No. 6500431el Inhibitors of Angiogenesis and Tumor Growth

FILE REFERENCE: 01/986-000410US

CURRENT APPLICATION NUMBER: US/09/352,548

CURRENT FILING DATE: 1999-07-12

EARLIER APPLICATION NUMBER: US 60/092,647

EARLIER FILING DATE: 1998-07-13

NUMBER OF SEQ ID NOS: 59

SOFTWARE: Patent In Ver. 2.1

SEQ ID NO 2

LENGTH: 81

TYPE: PRT

ORGANISM: Homo sapiens

FEATURE:

OTHER INFORMATION: Saposin B

US-09-352-548-2

Query Match 15.5%; Score 431; DB 4; Length 81;

Best Local Similarity 100.0%; Pred. No. 1.7e-31;

Matches 81; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 195 GDVCCDICIQWVTDIQTAVRTNSTFVQALVEHVEKCDRLGPMADICKNYISQYSEIAIQ 254  
DB 1 GDVCCDICIQWVTDIQTAVRTNSTFVQALVEHVEKCDRLGPMADICKNYISQYSEIAIQ 60

OY 255 MAMHOPKEICLVGFCDEVK 275  
DB 61 MAMHOPKEICLVGFCDEVK 81

## RESULT 9

US-08-584-671-15  
Sequence 15, Application US/08584671  
Patent No. 5910568

## GENERAL INFORMATION:

APPLICANT: HAMMERSTEDT, ROY H, BARBATO, GUY F,  
TITLE OF INVENTION: MOLECULE INVOLVED IN BINDING OF SPERM  
TO EGG SURFACES AND PROCEDURES FOR USE OF THIS MOLECULE  
NUMBER OF SEQUENCES: 16  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: INTELLECTUAL PROPERTY OFFICE, THE PENNSYLVANIA  
STATE UNIVERSITY  
STREET: 113 TECHNOLOGY CENTER  
CITY: UNIVERSITY PARK  
STATE: PENNSYLVANIA  
COUNTRY: UNITED STATES OF AMERICA  
ZIP: 16802-7000

COMPUTER READABLE FORM:  
MEDIUM TYPE: FLOPPY DISK  
COMPUTER: NEC 286  
OPERATING SYSTEM: DOS  
SOFTWARE: WORDPERFECT 5.1  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/584,671  
FILING DATE:  
CLASSIFICATION: 435

## ATTORNEY/AGENT INFORMATION:

NAME: MONAHAN, THOMAS J  
REGISTRATION NUMBER: 29835  
REFERENCE/DOCKET NUMBER:  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 814-865-6277  
TELEFAX: 814-865-3591  
INFORMATION FOR SEQ ID NO: 15:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 80  
TYPE: AMINO ACID  
STRANDEDNESS: SINGLE  
TOPOLOGY: UNKNOWN

US-08-584-671-15

Query Match 15.0%; Score 418.5; DB 2; Length 80;  
Best Local Similarity 98.8%; Pred. No. 2.2e-30;  
Matches 80; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

OY 128 PGEVCSALNLCESLQKHLAEINHOQLESNKIPBLDMTEVVAFPMANIPLLIYPQDGP 187  
DB 1 PGEVCSALNLCESLQKHLAEINHOQLESNKIPBLDMTEVVAFPMANIPLLIYPQDGP 187  
OY 188 KPQKNDGVCDCCIOGMTDI 208  
DB 60 KPQKNDGVCDCCIOGMTDI 80

## RESULT 10

US-09-027-376-15  
Sequence 15, Application US/09027376  
Patent No. 6004586

## GENERAL INFORMATION:

APPLICANT: HAMMERSTEDT, ROY H, BARBATO, GUY F,  
TITLE OF INVENTION: MOLECULE INVOLVED IN BINDING OF SPERM  
TO EGG SURFACES AND PROCEDURES FOR USE OF THIS MOLECULE  
NUMBER OF SEQUENCES: 16  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: INTELLECTUAL PROPERTY OFFICE, THE PENNSYLVANIA  
STATE UNIVERSITY  
STREET: 113 TECHNOLOGY CENTER  
CITY: UNIVERSITY PARK  
STATE: PENNSYLVANIA  
COUNTRY: UNITED STATES OF AMERICA  
ZIP: 16802-7000  
COMPUTER READABLE FORM:  
MEDIUM TYPE: FLOPPY DISK  
COMPUTER: NEC 286  
OPERATING SYSTEM: DOS  
SOFTWARE: WORDPERFECT 5.1  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/027,376  
FILING DATE:  
CLASSIFICATION: 435

## CORRESPONDENCE ADDRESS:

ADDRESSEE: INTELLECTUAL PROPERTY OFFICE, THE PENNSYLVANIA  
STATE UNIVERSITY  
STREET: 113 TECHNOLOGY CENTER  
CITY: UNIVERSITY PARK  
STATE: PENNSYLVANIA  
COUNTRY: UNITED STATES OF AMERICA  
ZIP: 16802-7000  
COMPUTER READABLE FORM:  
MEDIUM TYPE: FLOPPY DISK  
COMPUTER: NEC 286  
OPERATING SYSTEM: DOS  
SOFTWARE: WORDPERFECT 5.1  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/027,376  
FILING DATE:  
CLASSIFICATION: 435

PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/584,671  
FILING DATE:

## ATTORNEY/AGENT INFORMATION:

NAME: MONAHAN, THOMAS J  
REGISTRATION NUMBER: 29835  
REFERENCE/DOCKET NUMBER:  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 814-865-6277  
TELEFAX: 814-865-3591  
INFORMATION FOR SEQ ID NO: 15:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 80  
TYPE: AMINO ACID  
STRANDEDNESS: SINGLE  
TOPOLOGY: UNKNOWN

US-09-027-376-15

Query Match 15.0%; Score 418.5; DB 3; Length 80;  
Best Local Similarity 98.8%; Pred. No. 2.2e-30;  
Matches 80; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

OY 128 PGEVCSALNLCESLQKHLAEINHOQLESNKIPBLDMTEVVAFPMANIPLLIYPQDGP 187  
DB 1 PGEVCSALNLCESLQKHLAEINHOQLESNKIPBLDMTEVVAFPMANIPLLIYPQDGP 187  
OY 188 KPQKNDGVCDCCIOGMTDI 208  
DB 60 KPQKNDGVCDCCIOGMTDI 80

## RESULT 11

US-09-094-192-15  
Sequence 15, Application US/09094192  
Patent No. 6103483

## GENERAL INFORMATION:

APPLICANT: HAMMERSTEDT, ROY H, BARBATO, GUY F,  
TITLE OF INVENTION: MOLECULE INVOLVED IN BINDING OF SPERM TO EGG SURFACES AND PROCI  
NUMBER OF SEQUENCES: 16  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: INTELLECTUAL PROPERTY OFFICE, THE PENNSYLVANIA STATE UNIVERSITY  
STREET: 113 TECHNOLOGY CENTER  
CITY: UNIVERSITY PARK  
STATE: PENNSYLVANIA  
COUNTRY: UNITED STATES OF AMERICA  
ZIP: 16802-7000  
COMPUTER READABLE FORM:  
MEDIUM TYPE: FLOPPY DISK  
COMPUTER: NEC 286  
OPERATING SYSTEM: DOS  
SOFTWARE: WORDPERFECT 5.1  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/094,192  
FILING DATE:  
CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:  
NAME: MONAHAN, THOMAS J  
REGISTRATION NUMBER: 29835  
REFERENCE/DOCKET NUMBER:  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 814-865-6277  
TELEFAX: 814-865-3591  
INFORMATION FOR SEQ ID NO: 15:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 80  
TYPE: AMINO ACID  
STRANDEDNESS: SINGLE  
TOPOLOGY: UNKNOWN  
US-09-094-192-15

Query Match 14.9%; Score 415.5; DB 3; Length 80;  
Best Local Similarity 97.5%; Pred. No. 4.1e-30;  
Matches 79; Conservative 1; Mismatches 0; Indels 1; Gaps 1;

Qy 128 PGEVSAINTLCSLQKHIAELNHQKLESNKIPELDMTEVVAPEFMANIPILLYPODGPBS 187  
Db 1 PGEVC-AINTLCSLQKHIAELNHQKLESNKIPELDMTEVVAPEFMANIPILLYPODGPBS 59

Qy 188 KPOPKNGDVCCDCIOWTDTI 208  
Db 60 KPOPKNGDVCCDCIOWTDTI 80

RESULT 12  
US-08-100-247-3  
Sequence 3, Application US/08100247  
Patent No. 5571787  
GENERAL INFORMATION:  
APPLICANT: O'BRIEN, JOHN S.  
APPLICANT: KISHIMOTO, YASUO  
TITLE OF INVENTION: PROSAPOSIN AS A NEUROTROPHIC FACTOR  
NUMBER OF SEQUENCES: 5  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: KNOBBE, MARTENS, OLSON AND BEAR  
STREET: 620 NEWPORT CENTER DRIVE SIXTEENTH FLOOR  
CITY: NEWPORT BEACH  
STATE: CA  
COUNTRY: USA  
ZIP: 92660  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
OPERATING SYSTEM: IBM PC compatible  
SOFTWARE: Patent In Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/100.247  
FILING DATE: 19930730  
CLASSIFICATION: 514  
ATTORNEY/AGENT INFORMATION:  
NAME: Israelien, Ned A.  
REGISTRATION NUMBER: 29,655  
REFERENCE/DOCKET NUMBER: OBTEIN.002A  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 619-235-8550  
TELEFAX: 619-235-0176  
INFORMATION FOR SEQ ID NO: 3:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 80 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
HYPOTHETICAL: NO  
ANTI-SENSE: NO  
FRAGMENT TYPE: internal  
IMMEDIATE SOURCE:  
CLONE: SAPOSIN C  
US-08-100-247-3

Query Match 14.8%; Score 412; DB 1; Length 80;  
Best Local Similarity 100.0%; Pred. No. 8.5e-30;  
Matches 80; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 311 SDVYCEVEFLVKEVTKLIDNNKTEKELIDAFDMCKSLPKSLSECCQVVDYGSSTLS 370  
Db 1 SDVYCEVEFLVKEVTKLIDNNKTEKELIDAFDMCKSLPKSLSECCQVVDYGSSTLS 60

Qy 371 ILLEEVSPELVCSMLHLCSG 390  
Db 61 ILLEEVSPELVCSMLHLCSG 80

RESULT 13  
US-08-483-146A-3  
Sequence 3, Application US/08483146A  
Patent No. 5696080  
GENERAL INFORMATION:  
APPLICANT: O'Brien, John S.  
APPLICANT: Kishimoto, Yasuo  
TITLE OF INVENTION: PHARMACEUTICAL COMPOSITIONS  
TITLE OF INVENTION: COMPRISING PROSAPOSIN AND NEUROTROPHIC PEPTIDES DERIVED  
FROM  
NUMBER OF SEQUENCES: 11  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Knobbe, Martens, Olson and Bear  
STREET: 620 Newport Center Blvd. 16th Floor  
CITY: Newport Beach  
STATE: CA  
COUNTRY: USA  
ZIP: 92660  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
OPERATING SYSTEM: DOS  
SOFTWARE: FASTSEQ for Windows Version 2.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/483.146A  
FILING DATE: 07-JUN-1995  
CLASSIFICATION: 514  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER:  
FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: Israelien, Ned A.  
REGISTRATION NUMBER: 29,655  
REFERENCE/DOCKET NUMBER: MYELOS.002DV1  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 619-235-8550  
TELEFAX: 619-235-0176  
INFORMATION FOR SEQ ID NO: 3:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 80 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
FRAGMENT TYPE: internal  
US-08-483-146A-3

Query Match 14.8%; Score 412; DB 1; Length 80;  
Best Local Similarity 100.0%; Pred. No. 8.5e-30;  
Matches 80; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 311 SDVYCEVEFLVKEVTKLIDNNKTEKELIDAFDMCKSLPKSLSECCQVVDYGSSTLS 370  
Db 1 SDVYCEVEFLVKEVTKLIDNNKTEKELIDAFDMCKSLPKSLSECCQVVDYGSSTLS 60

Qy 371 ILLEEVSPELVCSMLHLCSG 390  
Db 61 ILLEEVSPELVCSMLHLCSG 80

## RESULT 14

US-08-232-513A-4

Sequence 4, Application US/08232513A  
Patent No. 5700909

GENERAL INFORMATION:

APPLICANT: O'Brien, John S.

TITLE OF INVENTION: Prosaposin and Cytokine-Derived Peptides

TITLE OF INVENTION: as Therapeutic Agents

NUMBER OF SEQUENCES: 20

CORRESPONDENCE ADDRESS:

ADDRESSEE: Campbell &amp; Flores LLP

STREET: 4370 La Jolla Village Drive, Suite 700

CITY: San Diego

STATE: California

COUNTRY: USA

ZIP: 92122

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/232,513A

FILING DATE: 21-Apr-1994

CLASSIFICATION: 514

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/100,247

FILING DATE: 30-JUL-1993

ATTORNEY/AGENT INFORMATION:

NAME: Campbell, Cathryn A.

REGISTRATION NUMBER: 31,815

REFERENCE/DOCKET NUMBER: P-TUD 1643

TELECOMMUNICATION INFORMATION:

TELEPHONE: (619) 535-9001

TELEFAX: (619) 535-8949

INFORMATION FOR SEQ ID NO: 4:

SEQUENCE CHARACTERISTICS:

LENGTH: 80 amino acids

TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: peptide

FEATURE:

NAME/KEY: Peptide

LOCATION: 1..80

OTHER INFORMATION: /label= Saposin\_C

US-08-232-513A-4

Query Match 14.8%; Score 412; DB 1; Length 80;

Best Local Similarity 100.0%; Pred. No. 8.5e-30;

Matches 80; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 311 SDVYCEVCEFLVKEVTKLIDNNKTEKEIIDAFDKMCKSLPKSLSEBCOEVDVTGSSILS 370

Db 1 SDVYCEVCEFLVKEVTKLIDNNKTEKEIIDAFDKMCKSLPKSLSEBCOEVDVTGSSILS 60

OY 371 ILLEEVSPELVCSMLHLCG 390

Db 61 ILLEEVSPELVCSMLHLCG 80

OY 61 ILLEEVSPELVCSMLHLCG 80

Db 61 ILLEEVSPELVCSMLHLCG 80

## RESULT 15

US-08-484-594A-3

Sequence 3, Application US/08484594A  
Patent No. 5714459

GENERAL INFORMATION:

APPLICANT: O'Brien, John S.

APPLICANT: Kishimoto, Yasuo

TITLE OF INVENTION: USE OF PROSAPOSIN AND NEUROTROPHIC PEPTIDES

TITLE OF INVENTION: DERIVED THEREFROM

NUMBER OF SEQUENCES: 11

CORRESPONDENCE ADDRESS:

ADDRESSEE: Knobbe, Martens, Olson and Bear

STREET: 620 Newport Center Drive, Sixteenth Floor

CITY: Newport Beach

STATE: CA

COUNTRY: USA

ZIP: 92660

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette

COMPUTER: IBM Compatible

OPERATING SYSTEM: DOS

SOFTWARE: PastSEO for Windows Version 2.0

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/484,594A

FILING DATE: 07-JUN-1995

CLASSIFICATION: 514

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 08/100,247

FILING DATE: 30-JUL-1993

ATTORNEY/AGENT INFORMATION:

NAME: Israelien, Ned A.

REGISTRATION NUMBER: 29,655

REFERENCE/DOCKET NUMBER: MYELOS.002DV2

TELECOMMUNICATION INFORMATION:

TELEPHONE: 619-235-8550

TELEFAX: 619-235-0176

INFORMATION FOR SEQ ID NO: 3:

SEQUENCE CHARACTERISTICS:

LENGTH: 80 amino acids

TYPE: amino acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: protein

FRAGMENT TYPE: internal

US-08-484-594A-3

Query Match 14.8%; Score 412; DB 1; Length 80;

Best Local Similarity 100.0%; Pred. No. 8.5e-30;

Matches 80; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 311 SDVYCEVCEFLVKEVTKLIDNNKTEKEIIDAFDKMCKSLPKSLSEBCOEVDVTGSSILS 370

Db 1 SDVYCEVCEFLVKEVTKLIDNNKTEKEIIDAFDKMCKSLPKSLSEBCOEVDVTGSSILS 60

OY 371 ILLEEVSPELVCSMLHLCG 390

Db 61 ILLEEVSPELVCSMLHLCG 80

OY 61 ILLEEVSPELVCSMLHLCG 80

Db 61 ILLEEVSPELVCSMLHLCG 80

Search completed: May 5, 2004, 13:16:43  
Job time : 23.8908 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

## OM protein - protein search, using sw model

Run on: May 5, 2004, 13:15:48 ; Search time 73.151 Seconds

(without alignments)  
2023.963 Million cell updates/sec

Title: US-09-743-684A-1

Sequence: 1 MYALFLIASLGAALAGPVL.....NTETAAQCNVHECKRHWVN 524

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters: 1586107

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A\_Geneseq\_29Jan04:\*

1: geneseqp1980s:\*  
2: geneseqp1990s:\*  
3: geneseqp2000s:\*  
4: geneseqp2001s:\*  
5: geneseqp2002s:\*  
6: geneseqp2003as:\*  
7: geneseqp2003bs:\*  
8: geneseqp2004s:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2789	100.0	524	2	AAR70783
2	2789	100.0	524	2	AAW85652
3	2789	100.0	524	3	AAV58716
4	2789	100.0	524	6	ABU79099
5	2789	100.0	524	6	ABU05200
6	2789	100.0	524	6	ABU05207
7	2789	100.0	524	6	ABU05203
8	2789	100.0	524	6	ABU07340
9	2789	100.0	524	6	ABU05216
10	2789	100.0	524	6	ABU05202
11	2789	100.0	524	6	ABU05214
12	2789	100.0	524	6	ABU05215
13	2789	100.0	524	6	ABU05199
14	2789	100.0	524	6	ABU05212
15	2789	100.0	524	6	ABU05213
16	2789	100.0	524	6	ABU05205
17	2779	99.6	527	6	ABU05208
18	2777.5	99.6	527	5	ABP68602
19	2777.5	99.6	527	5	ABU79100
20	2777.5	99.6	527	6	ABU05204
21	2777.5	99.6	527	6	ABU05210
22	2772.5	99.4	523	4	AAH31916
23	2772.5	99.4	523	6	ABU05211
24	2768	99.2	526	6	ABU05209

26	2767.5	99.2	527	6	ABU05206
27	2449.5	87.8	479	6	ABR39442
28	2024.5	72.6	385	6	ABR41750
29	1969	70.6	354	7	ADB85295
30	1921	68.9	554	5	ABR57102
31	1201	43.1	521	3	AAV56966
32	1201	43.1	521	4	AAW78587
33	1201	43.1	521	4	AAW79571
34	1201	43.1	521	4	AAW86362
35	1201	43.1	521	4	AAE15547
36	1196	42.9	521	4	AAW79572
37	1139.5	40.9	531	5	AAE01770
38	1139.5	40.9	531	5	ABG64156
39	1125	40.3	210	6	ABU05201
40	1093	39.2	209	5	ABG70166
41	834.5	29.9	362	4	AAW78588
42	834.5	29.9	362	5	ABW07499
43	812	29.1	153	6	ABU70504
44	811	29.1	153	6	ABU70799
45	676	24.2	129	6	ABU70422

## ALIGNMENTS

RESULT 1  
AAR70783  
ID AAR70783 standard; protein; 524 AA.  
XX  
AC AAR70783;  
XX  
DT 25-MAR-2003 (revised)  
DT 30-AUG-1995 (first entry)  
XX  
DE Prosapoin.  
XX  
KW Saposin-C; neuron; myelination; nervous system; neuroblastoma;  
KW neurotrophic peptide; multiple sclerosis; leukoencephalitis;  
KW adrenal leukodystrophy; prosapoin.  
XX  
OS Homo sapiens.  
XX  
PN W09503821-A1.  
XX  
PD 09-FEB-1995.  
XX  
PF 28-JUL-1994; 94WO-US008453.  
XX  
PR 30-JUL-1993; 93US-00100247.  
PR 21-APR-1994; 94US-00232513.  
XX  
PA (OBRI/) O'BRIEN J S.  
PI O'Brien JS, Kishimoto Y,  
XX WPI, 1995-082029/11.  
XX DR N-PSDB; AAQ85355.  
XX  
PT Stimulating neural cell out-growth and myelination - with pro:saposin,  
PT saposin C or new neurotrophic peptide(s) from cytokine(s), for treating  
PT nervous system diseases.  
XX  
PS Disclosure; Page 30-32; 50pp; English.  
XX  
XX The peptide given in AAR70773, corresponding to amino acids 8-29 of human  
XX saposin-C (AAR70784), promotes neurite outgrowth in vitro. A consensus  
XX sequence was determined by comparing the peptide with hematopoietic and  
XX neurotrophic cytokines, and neurotrophic peptides (AAR70774-82) were  
XX identified in the AB loop of human ciliary neurotrophic factor,  
XX interleukins-6, -2, -3 and -gamma, erythropoietin and leukocyte  
XX inhibitory factor, and in helix C of human interleukin-1-beta and  
XX oncostatin-M. Prosapoin (AAR70783) and saposin-C also promoted nerve  
XX cell myelination ex vivo. (Updated on 25-MAR-2003 to correct PN field.)

CC (Updated on 25-MAR-2003 to correct PA field.) (Updated on 25-MAR-2003 to  
 CC correct PI field.)

CC Sequence 524 AA;

Query Match 100.0%; Score 2789; DB 2; Length 524;  
 Best Local Similarity 100.0%; Pred. No. 2,3e-237;  
 Matches 524; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY 1 MYALFLASLILGALLAGPVILGLKECTRGSAVWCQNVKTASDCAVGHCLQTVWKNKPTVKS 60
DB 1 MYALFLASLILGALLAGPVILGLKECTRGSAVWCQNVKTASDCAVGHCLQTVWKNKPTVKS 60
QY 61 LPEDICKDVVTAAGDMLKDNATEEELIYVLEKTCMDLPKPNMSASCKEIVDSYLVILDI 120
DB 61 LPEDICKDVVTAAGDMLKDNATEEELIYVLEKTCMDLPKPNMSASCKEIVDSYLVILDI 120
QY 121 IGENSRPGEVCSALNLCESLQKHLAELNHQKLESNNKIPBLDMEVVAPEMANIPILLY 180
DB 121 IGENSRPGEVCSALNLCESLQKHLAELNHQKLESNNKIPBLDMEVVAPEMANIPILLY 180
QY 181 PDGGRSKRPQKNDGVQDCIQWVTDIQTAVRTNSTFVQALVHVKECDRLGPGMADI 240
DB 181 PDGGRSKRPQKNDGVQDCIQWVTDIQTAVRTNSTFVQALVHVKECDRLGPGMADI 240
QY 241 CKNYISQYSEIALQMMHMQPEICALVGFCEVKEPMQTLVPAKVASKNVIPLELIVE 300
DB 241 CKNYISQYSEIALQMMHMQPEICALVGFCEVKEPMQTLVPAKVASKNVIPLELIVE 300
QY 301 PIKKEHVPAKSDVYVEVEEFLVKEVTYKILDNKTEKEILDAFDKXCSKLPKSLSECEV 360
DB 301 PIKKEHVPAKSDVYVEVEEFLVKEVTYKILDNKTEKEILDAFDKXCSKLPKSLSECEV 360
QY 361 VDTYSSILSLILEVSEPELVCSMLHCSGTRLPALTVHVOQPDGGECEVCKLVGLD 420
DB 361 VDTYSSILSLILEVSEPELVCSMLHCSGTRLPALTVHVOQPDGGECEVCKLVGLD 420
QY 421 RNLEKNSTKQEBILALEKGSFLLPDPYQKCDQFAVEYEPILIELVEMPSFVCLXIG 480
DB 421 RNLEKNSTKQEBILALEKGSFLLPDPYQKCDQFAVEYEPILIELVEMPSFVCLXIG 480
QY 481 ACPSAHKLPLGTEKICMGPSWCONTEPAACNAVEHCKRVWN 524
DB 481 ACPSAHKLPLGTEKICMGPSWCONTEPAACNAVEHCKRVWN 524

```

RESULT 2  
 AAW85652 standard; protein; 524 AA.

XX AAW85652;

XX 19-JUL-1999 (first entry)

XX Human prosaposin N-terminal peptide.

XX Prosaposin; saposin; prosapetides; prosaposin receptor agonists; PRA;  
 XX peripheral nervous system; central nervous system; PNS; CNS; Akt; Bcl-2;  
 XX therapy; treatment; apoptosis; caspase; tumor necrosis factor; TNF;  
 XX cytokine; interferon gamma; IFN; inflammation; rheumatoid arthritis;  
 XX Crohn's disease; irritable bowel syndrome; ashma; cardiac infarction;  
 XX congestive heart failure; multiple sclerosis;  
 XX acute disseminated inflammatory leukoencephalitis;  
 XX progressive multifocal leukoencephalitis; Alzheimer's disease;  
 XX Parkinson's disease; amyotrophic lateral sclerosis; Huntington's disease;  
 XX ischemic heart disease; Guillain-Barre disease; alopecia; AIDS dementia;  
 XX cerebral malaria; HTLV; neuropathy;  
 XX inflammatory neurodegenerative disease; toxin-induced liver disease.

OS Homo sapiens.

XX MO9912559-A1.

PD 18-MAR-1999.

XX 09-SEP-1998; 98WO-US019216.

XX 09-SEP-1997; 97US-0058352P.

XX 04-JUN-1998; 98US-0088129P.

XX (REGC) UNIV CALIFORNIA.

XX O'Brien US;

XX WPI; 1999-228139/19.

XX N-PSDB; AAX08488.

XX Use of prosaposin receptor agonist.

XX Claim 7; Fig 2; 90pp; English.

Prosaposin is a 70kDa glycoprotein which is proteolytically processed to generate saposins A, B, C and D, all of which are similar to each other and have a similar placement of six cysteines, a glycosylation site and conserved proline residues. Prosaposin, saposin C and prosaposin derived peptides (prosapetides), have therapeutic applications in promoting recovery after toxic, traumatic, myocardial ischemic, degenerative and inherited lesions to the peripheral and central nervous system. Prosaposin receptor agonists (PRAs) inhibit proinflammatory cytokine-induced apoptosis by activation of the Ser/Thr protein kinase Akt. Akt releases Bcl-2 and its family members which inhibit caspases, thereby inhibiting apoptosis. An additional mechanism whereby PRAs inhibit apoptosis is by blocking activation of JNK, a proapoptotic signaling component. Within several minutes after binding to the receptor, PRAs block JNK activation induced by tumor necrosis factor- $\alpha$  (TNF  $\alpha$ ). The activation of JNK by TNF  $\alpha$  is another well known mechanism for TNF  $\alpha$ -induced, as well as other proinflammatory cytokine-induced apoptosis. The method can be used for inhibiting apoptosis which is caspase-mediated or induced by a proinflammatory cytokine which is TNF  $\alpha$  or interferon- $\gamma$ . It can be used for inhibiting apoptosis associated with a disorder such as e.g. rheumatoid arthritis, Crohn's disease, irritable bowel syndrome, asthma, cardiac infarction, congestive heart failure, multiple sclerosis, acute disseminated inflammatory leukoencephalitis, progressive multifocal leukoencephalitis, Alzheimer's disease, Parkinson's disease, amyotrophic lateral sclerosis, Huntington's disease, ischemic heart disease, Guillain-Barre disease, traumatic brain injury, traumatic spinal cord injury, alopecia, AIDS dementia, cerebral malaria, HTLV, neuropathy, inflammatory neurodegenerative disease, and toxin-induced liver disease. This 524 N-terminal peptide of prosaposin also acts as a prosaposin receptor agonist

XX Sequence 524 AA;

Query Match 100.0%; Score 2789; DB 2; Length 524;

Best Local Similarity 100.0%; Pred. No. 2,3e-237;  
 Matches 524; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY 1 MYALFLASLILGALLAGPVILGLKECTRGSAVWCQNVKTASDCAVGHCLQTVWKNKPTVKS 60
DB 1 MYALFLASLILGALLAGPVILGLKECTRGSAVWCQNVKTASDCAVGHCLQTVWKNKPTVKS 60
QY 61 LPEDICKDVVTAAGDMLKDNATEEELIYVLEKTCMDLPKPNMSASCKEIVDSYLVILDI 120
DB 61 LPEDICKDVVTAAGDMLKDNATEEELIYVLEKTCMDLPKPNMSASCKEIVDSYLVILDI 120
QY 121 IGENSRPGEVCSALNLCESLQKHLAELNHQKLESNNKIPBLDMEVVAPEMANIPILLY 180
DB 121 IGENSRPGEVCSALNLCESLQKHLAELNHQKLESNNKIPBLDMEVVAPEMANIPILLY 180
QY 181 PDGGRSKRPQKNDGVQDCIQWVTDIQTAVRTNSTFVQALVHVKECDRLGPGMADI 240
DB 181 PDGGRSKRPQKNDGVQDCIQWVTDIQTAVRTNSTFVQALVHVKECDRLGPGMADI 240
QY 241 CKNYISQYSEIALQMMHMQPEICALVGFCEVKEPMQTLVPAKVASKNVIPLELIVE 300
DB 241 CKNYISQYSEIALQMMHMQPEICALVGFCEVKEPMQTLVPAKVASKNVIPLELIVE 300

```



```

Cc 241 CKNYISQYSEIAIQMMHMQPKEICALVGFCDVEKEMPMQTLVPAKASKNVIPLALVE 300
Cc 301 PIKHEVPKSDVYCEVCEFLVKEVTKLIDNNKTEKEILDAPDKMCSLPSLSSECOEV 360
Cc 301 PIKHEVPKSDVYCEVCEFLVKEVTKLIDNNKTEKEILDAPDKMCSLPSLSSECOEV 360
Cc 361 VDTYSSILSLILEVSPBELVCSMLHLCSGTRLPALTHTVHTQPKDGGCEVCKLVGLD 420
Cc 361 VDTYSSILSLILEVSPBELVCSMLHLCSGTRLPALTHTVHTQPKDGGCEVCKLVGLD 420
Cc 421 RNLEKSTKQEIILALEKGSFLPDPYQKQCDQFAVEYEPVLIIELVMPSPFVCLKIG 480
Cc 421 RNLEKSTKQEIILALEKGSFLPDPYQKQCDQFAVEYEPVLIIELVMPSPFVCLKIG 480
Cc 481 ACPSAHKPLGTGTEKCIWGPSYWCNTETAAOCNAVEHCKRHVMN 524
Cc 481 ACPSAHKPLGTGTEKCIWGPSYWCNTETAAOCNAVEHCKRHVMN 524

RESULT 3
AAVS8716
ID AAVS8716 standard; protein, 524 AA.
AC AAVS8716;
DT 25-APR-2000 (first entry)
DE Human prosaposin.
XX
XX Prosapoin; saposin B; antiangiogenic; angiogenesis inhibitor;
XX antitumor; antiproliferative; antimitratory; Kaposi's sarcoma; tumour;
XX human; therapy.
XX
XX Homo sapiens.
XX
XX Key Location/Qualifiers
XX Protein 195..275
XX /note="mature saposin B"
XX Peptide 195..205
XX /note="specifically claimed antiangiogenic peptide of
XX Peptide 196..200
XX /note="specifically claimed antiangiogenic peptide of
XX Claim 4"
XX
XX WO200002902-A1.
XX
XX 20-JAN-2000.
XX
XX 12-JUL-1999; 99WO-US015772.
XX
XX 13-JUL-1998; 98US-0092647P.
XX
XX (GILL/) GILL P S.
XX
XX GILL PS;
XX
XX WPI; 2000-171128/15.
XX
XX Saposin B derived peptides, useful as inhibitors of angiogenesis and
XX tumor growth.
XX
XX Disclosure; Page 18; 78pp; English.
XX
XX The present sequence is that of human prosaposin, a precursor of saposin
XX B. The invention is based on the discovery that saposin B, previously
XX known to be involved in the hydrolysis of sphingolipids, has potent
XX antiangiogenic and antitumor activity, and also has antiproliferative
XX and antimigratory activity against endothelial cells. This activity is
XX conserved in cryptic polypeptides as small as 5 amino acids (see AAVS8684
XX -715), which can be synthetically prepared and used in vitro or in vivo
XX for the treatment of undesired angiogenesis and tumor growth, especially
XX Kaposi's sarcoma (claimed). The polypeptides can also be used in

```

```

Cc conjunction with cytotoxic moieties to selectively kill certain cell
Cc types, e.g. for treatment of cancer, angiofibroma, neovascular glaucoma,
Cc arteriovenous malformation, nonunion fracture, arthritis and other
Cc connective tissue disorders, Osler-Weber syndrome, atherosclerotic
Cc plaque, psoriasis, corneal graft neovascularization, pyogenic granuloma,
Cc retrorenal fibroplasia, diabetic retinopathy, scleroderma, haemangioma,
Cc trachoma, vascular adhesions and hypertrophic scars
Cc
Cc Sequence 524 AA;
Cc
Cc Query Match 100.0%; Score 2789; DB 3; Length 524;
Cc Best Local Similarity 100.0%; Pred. No. 2,3e-237;
Cc Matches 524; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Cc
Cc 1 MYALFLASLIGALAPVIGLKECTGSAWCONVKTASCGVAKCLQTVNKKPYKS 60
Cc 1 MYALFLASLIGALAPVIGLKECTGSAWCONVKTASCGVAKCLQTVNKKPYKS 60
Cc 61 LPDICKDVTAAGDMLKDNATEEELVYLEKTCDMILPKPMNSASCKEIVDSYPLIIDI 120
Cc 61 LPDICKDVTAAGDMLKDNATEEELVYLEKTCDMILPKPMNSASCKEIVDSYPLIIDI 120
Cc 61 LPDICKDVTAAGDMLKDNATEEELVYLEKTCDMILPKPMNSASCKEIVDSYPLIIDI 120
Cc 121 IKGMSRGEVCSALNLCESLQKHLAELNHQKQESNKPILDMTEVVAPEPMANIPLLLY 180
Cc 121 IKGMSRGEVCSALNLCESLQKHLAELNHQKQESNKPILDMTEVVAPEPMANIPLLLY 180
Cc 181 PDDGPRSKPQPKDNGDYCCDCTQWVTDIQTAVRNTSTFVQALVHVKEECRLGGMADI 240
Cc 181 PDDGPRSKPQPKDNGDYCCDCTQWVTDIQTAVRNTSTFVQALVHVKEECRLGGMADI 240
Cc 181 PDDGPRSKPQPKDNGDYCCDCTQWVTDIQTAVRNTSTFVQALVHVKEECRLGGMADI 240
Cc 241 CKNYISQYSEIAIQMMHMQPKEICALVGFCDVEKEMPMQTLVPAKASKNVIPLALVE 300
Cc 241 CKNYISQYSEIAIQMMHMQPKEICALVGFCDVEKEMPMQTLVPAKASKNVIPLALVE 300
Cc 241 CKNYISQYSEIAIQMMHMQPKEICALVGFCDVEKEMPMQTLVPAKASKNVIPLALVE 300
Cc 301 PIKHEVPKSDVYCEVCEFLVKEVTKLIDNNKTEKEILDAPDKMCSLPSLSSECOEV 360
Cc 301 PIKHEVPKSDVYCEVCEFLVKEVTKLIDNNKTEKEILDAPDKMCSLPSLSSECOEV 360
Cc 361 VDTYSSILSLILEVSPBELVCSMLHLCSGTRLPALTHTVHTQPKDGGCEVCKLVGLD 420
Cc 361 VDTYSSILSLILEVSPBELVCSMLHLCSGTRLPALTHTVHTQPKDGGCEVCKLVGLD 420
Cc 421 RNLEKSTKQEIILALEKGSFLPDPYQKQCDQFAVEYEPVLIIELVMPSPFVCLKIG 480
Cc 421 RNLEKSTKQEIILALEKGSFLPDPYQKQCDQFAVEYEPVLIIELVMPSPFVCLKIG 480
Cc 481 ACPSAHKPLGTGTEKCIWGPSYWCNTETAAOCNAVEHCKRHVMN 524
Cc 481 ACPSAHKPLGTGTEKCIWGPSYWCNTETAAOCNAVEHCKRHVMN 524

RESULT 4
ABU79099
ID ABU79099 standard; protein, 524 AA.
AC ABU79099;
DT 18-JUN-2003 (first entry)
DE Lip-TAA binding protein, Prosaposin.
XX
XX Superantigen; SAg; staphylococcal enterotoxin; tumour; cancer; apoptosis;
XX gene therapy; mammalian cell receptor; cytostatic;
XX tumour associated lipid; energy; T cell; antigen presenting cell; APC;
XX tumouricidal immunocyte; antitumour.
XX
XX Unidentified.
XX
XX US2002177551-A1.
XX
XX 28-NOV-2002.
XX
XX 30-MAY-2001; 2001US-00870759.

```

XX 31-MAY-2000; 2000US-0208128P.  
 PR (TERM/) Terman D S.  
 XX Terman DS;  
 XX MPI; 2003-361759/34.  
 DR A mammalian cell receptor, useful in the treatment of cancer by binding  
 XX to tumor associated lipids where the binding induces anergy or apoptosis  
 PT in T cells and antigen presenting cells.  
 XX Disclosure; Page; 167pp; English.  
 PS  
 XX The invention relates to a mammalian cell receptor, useful in the  
 CC treatment of cancer, which binds to tumour associated lipids and induces  
 CC anergy or apoptosis in the T cells and antigen presenting cells (APCs).  
 CC Also included are a mammalian cell useful in the treatment of cancer  
 CC where the receptor which binds tumour associated lipids and induces  
 CC cellular inactivation or death is deleted or functionally deactivated,  
 CC producing (M1) a tumouricidal immunocyte population in vivo in a mammal  
 CC (by allowing tumour associated fatty acids, ceramides, glycolipids,  
 CC sphingolipids, glycosphingolipids, phosphosphingolipids, gangliosides,  
 CC sialylated glycans, lipopeptides and proteoglycolipids are inactivated or  
 CC deleted), a construct useful in the treatment of cancer comprising a  
 CC superantigen (SAG) nucleotide inserted into a virus, a mammalian T cell  
 CC useful in the treatment of cancer (where an adaptor protein which  
 CC inhibits T cell activation by tumour associated antigens is deleted or  
 CC functionally deactivated), a composition useful in the treatment of  
 CC cancer (comprising a lipid raft conjugated to a superantigen), producing  
 CC (M2) a tumouricidal immunocyte population ex vivo in a mammal (by  
 CC allowing tumour associated lipids to contact immunocytes, in which  
 CC receptors for the lipids are inactivated or deleted to produce a  
 CC tumouricidal immunocyte population, and administering the tumouricidal  
 CC activated immunocytes to the host), producing (M3) a tumouricidal APC  
 CC population ex vivo in a mammal (by allowing a tumour associated lipid to  
 CC contact APCs, in which receptors for the tumour associated lipids are  
 CC inactivated or deleted to produce a tumouricidally activated population,  
 CC and administering APCs to the host), producing a tumouricidal T cell  
 CC population ex vivo in a mammal (by allowing a tumour associated lipid to  
 CC contact T cells, in which adaptor proteins, which inhibit T cell  
 CC activation by tumour associated antigens, are deleted or functionally  
 CC deactivated to produce a tumouricidal population of T cells, and  
 CC administering the tumouricidally activated T cells to the host, or  
 CC allowing a superantigen-lipid raft to contact T cells ex vivo, and  
 CC administering the tumouricidally activated T cells to the host, treating  
 CC (M5) cancer in a mammal (by administering a lipid binding molecule which  
 CC binds immunosuppressive tumour associated lipids in vivo), producing (M6)  
 CC a tumouricidal T cell population in vivo in a mammal (by allowing a  
 CC tumour associated antigen to contact immunocytes in which adaptor  
 CC proteins which inhibit T cell activation by tumour associated antigens  
 CC are deleted or functionally deactivated) and producing (M7) a  
 CC tumouricidal T cell population ex vivo in a mammal comprising allowing a  
 CC superantigen-lipid raft conjugate to contact immunocytes in vivo. The  
 CC receptors, methods and compositions are useful for treating cancers and  
 CC tumours. Bacterial superantigens are co-administered or administered as  
 CC fusion constructs with anti-tumour proteins or motifs identifying a tumour  
 CC antigen, which can be functionally deactivated in the method of the  
 CC invention. Note: The sequence data for this patent did not form part of  
 CC the printed specification, but was obtained in electronic format from the  
 CC US patent office website at  
 CC "http://data.uspto.gov/sequence.htm?DocID=20020177551"  
 CC  
 XX Sequence 524 AA:  
 SQ

Query Match 100.0%; Score 2789; DB 6; Length 524;  
 Best Local Similarity 100.0%; Pred. No. 2,3e-237;  
 Matches 524; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 1 MYALFLASLILGALAGPVGLKECTRGSAVWCQNVKTASDCAVHCHCQTWNKPTVKS 60

Db  
 1 MYALFLASLILGALAGPVGLKECTRGSAVWCQNVKTASDCAVHCHCQTWNKPTVKS 60  
 QY LPDLCIKDVVTAAGDMLKDNATEEELIVLEKTCMDLPRKNMSASCKEIVDSYLPVILDT 120  
 Db LPDLCIKDVVTAAGDMLKDNATEEELIVLEKTCMDLPRKNMSASCKEIVDSYLPVILDT 120  
 QY IKGEMSRPGEVCSALNLCESLQKHAEINOKLESNKIPELDMTEVFAVPMNIPILLY 180  
 Db IKGEMSRPGEVCSALNLCESLQKHAEINOKLESNKIPELDMTEVFAVPMNIPILLY 180  
 QY 181 FQDGRSRKPPKQNDVCDQDCIQVTDIQTAVRTNSTFVQALVEHYEBCDRLPGWADI 240  
 Db 181 FQDGRSRKPPKQNDVCDQDCIQVTDIQTAVRTNSTFVQALVEHYEBCDRLPGWADI 240  
 QY 241 CKNYISQYSSEIATQMMHMQPKKICLVGFCDEVKEMPTQTLVPAVASKNVI PALELYE 300  
 Db 241 CKNYISQYSSEIATQMMHMQPKKICLVGFCDEVKEMPTQTLVPAVASKNVI PALELYE 300  
 QY 301 PIKKEVPAKSDVYCEVCEFLVKEVTKLIDNNTEKEILDAFDKMSKLPKLSSECOEY 360  
 Db 301 PIKKEVPAKSDVYCEVCEFLVKEVTKLIDNNTEKEILDAFDKMSKLPKLSSECOEY 360  
 QY 361 VDTYSSILSLLEEVSPPELVCSMLHLCSTRLPALTVHTQPKDGFCVCKKLVGYLD 420  
 Db 361 VDTYSSILSLLEEVSPPELVCSMLHLCSTRLPALTVHTQPKDGFCVCKKLVGYLD 420  
 QY 421 RNLKSTKOEILALAEKGSFLPDYKQCDQFVAVEBPVLEILVEWMDSPVCLTKIG 480  
 Db 421 RNLKSTKOEILALAEKGSFLPDYKQCDQFVAVEBPVLEILVEWMDSPVCLTKIG 480  
 QY 481 ACSNAKPLIGTEKCIWGSYWCNTETAACQNAVEHCKRHVWN 524  
 Db 481 ACSNAKPLIGTEKCIWGSYWCNTETAACQNAVEHCKRHVWN 524

RESULT 5  
 ABOU05200 standard; protein; 524 AA.  
 XX ABOU05200;  
 AC 29-JAN-2003 (first entry)  
 DT XX  
 DE Human expressed protein tag (EPT) #1866.  
 KW Translational profiling; expressed protein tag; EPT; kinase; phosphatase;  
 KW protease; protease inhibitor; transporter; cytoskeletal protein;  
 KW receptor; transcription factor; cancer; MHC;  
 KW major histocompatibility complex; myeloma; colon cancer; gastric cancer;  
 KW adenocarcinoma; sarcoma; melanoma; lymphoma; leukemia.  
 OS Homo sapiens.  
 XX  
 PN WO200278524-A2.  
 PD 10-OCT-2002.  
 XX  
 PF 28-MAR-2002; 2002MO-US009671.  
 XX  
 PR 28-MAR-2001; 2001US-0279495P.  
 XX  
 PR 21-MAY-2001; 2001US-0292544P.  
 XX  
 PR 08-AUG-2001; 2001US-0310801P.  
 XX  
 PR 01-OCT-2001; 2001US-0326370P.  
 XX  
 PR 04-DEC-2001; 2001US-0336780P.  
 XX  
 PR 20-FEB-2002; 2002US-0358985P.  
 XX  
 PA (ZYCO-) ZYCOs INC.  
 XX  
 PI Chicx RM, Tomlinson AJ, Urban RG;  
 XX MPI; 2003-040607/03.  
 DR  
 XX

PT New polypeptides (e.g. kinases, phosphatases, proteases, transporters,  
PT cytoskeletal proteins, receptors or transcription factors), useful for  
PT treating cancer, e.g. colon cancer, gastric cancer, sarcoma, lymphoma or  
PT leukemia.

XX Example 2; SEQ ID NO 1866; 134pp; English.

CC The invention describes a purified polypeptide, which comprises a  
CC fragment of a kinase, phosphatase, protease, protease inhibitor,  
CC transporter, cytoskeletal protein, receptor or transcription factor. The  
CC polypeptide is useful as an immunogenic composition for eliciting in a  
CC mammal an immunogenic response directed against any of the purified  
CC polypeptide. The purified polypeptide, or the antibody that binds to this  
CC polypeptide, is useful for treating cancer. The polypeptide is also  
CC useful for identifying compounds that binds to a naturally processed  
CC class I or class II MHC-binding polypeptide. The polypeptides and  
CC polynucleotides are particularly useful for treating or preventing  
CC myeloma, colon cancer, gastric cancer, adenocarcinoma, sarcoma, melanoma,  
CC lymphoma or leukemia. These are also useful for screening agents for  
CC treating the above mentioned diseases. This sequence represents an  
CC expressed protein tag (EPT) isolated from human tissue for translational  
CC profiling. Note: This sequence does not appear in the printed  
CC specification but was obtained in electronic format directly from WIPO at  
CC ftp.wipo.int/pub/published\_pct\_sequences

XX Sequence 524 AA;

Query Match 100.0%; Score 2789; DB 6; Length 524;  
Best Local Similarity 100.0%; Pred. No. 2,3e-237;  
Matches 524; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MYALFLASLLGALAGVILGLKECTRGSAVWCNVKTASDCGAVKHCLQTVWKNKPTVKS 60  
DB 1 MYALFLASLLGALAGVILGLKECTRGSAVWCNVKTASDCGAVKHCLQTVWKNKPTVKS 60  
QY 61 LPDCICKDVTAAGDMLKDNATEEELIYLEKTCQMLPKPMNSASCKEIVSYLPVILI 120  
DB 61 LPDCICKDVTAAGDMLKDNATEEELIYLEKTCQMLPKPMNSASCKEIVSYLPVILI 120  
QY 121 IKGEMSRGEVCSALNLCESLQKHIAELNHOKLESNKIPELDMTEVVAPEMANIPLLLY 180  
DB 121 IKGEMSRGEVCSALNLCESLQKHIAELNHOKLESNKIPELDMTEVVAPEMANIPLLLY 180  
QY 181 PODGRSRKPRKNDNDVQDCIQWVTDIQTAVRTNSTVQALVEHVKECCRLGGMADI 240  
DB 181 PODGRSRKPRKNDNDVQDCIQWVTDIQTAVRTNSTVQALVEHVKECCRLGGMADI 240  
QY 241 CKNYISQYSEIAIQMMHMQPKICALVGFCEVEMPMQTLVPAKVASKVITPALLEVE 300  
DB 241 CKNYISQYSEIAIQMMHMQPKICALVGFCEVEMPMQTLVPAKVASKVITPALLEVE 300  
QY 301 PIKHEVPAKSDVYCEVCEFLVKEVTKLIDNNKTEKEILDAFKMCSKLPKSLSECEV 360  
DB 301 PIKHEVPAKSDVYCEVCEFLVKEVTKLIDNNKTEKEILDAFKMCSKLPKSLSECEV 360  
QY 361 VDTYGSSTLILILEVSVBELVCSMTLCSGTRLPALTYHVTQPKGGGCEVCKLVGILD 420  
DB 361 VDTYGSSTLILILEVSVBELVCSMTLCSGTRLPALTYHVTQPKGGGCEVCKLVGILD 420  
QY 421 RNLEKNSTKQELIALLEKGCSEFLPDYQKCDQFAVEEPVILILEVEMDSFVCLKIG 480  
DB 421 RNLEKNSTKQELIALLEKGCSEFLPDYQKCDQFAVEEPVILILEVEMDSFVCLKIG 480  
QY 481 ACPSAHKPLGTGTEKCIWGPSYWCNTETTAQCNAAVEHCKRHVN 524  
DB 481 ACPSAHKPLGTGTEKCIWGPSYWCNTETTAQCNAAVEHCKRHVN 524

RESULT 6  
ABU05207  
ID ABU05207 standard; protein; 524 AA.  
XX  
AC ABU05207;

XX 29-JAN-2003 (first entry)  
DT  
XX  
XX Human expressed protein tag (EPT) #1873.  
DE  
XX  
XX Translational profiling; expressed protein tag; EPT; Kinase; phosphatase;  
KW  
KW protease; protease inhibitor; transporter; cytoskeletal protein;  
KW receptor; transcription factor; cancer; MHC;  
KW major histocompatibility complex; myeloma; colon cancer; gastric cancer;  
KW adenocarcinoma; sarcoma; melanoma; lymphoma; leukemia.  
XX  
XX Homo sapiens.  
OS  
XX  
XX WO200278524-A2.  
PD  
XX  
XX 10-OCT-2002.  
PD  
XX  
XX 28-MAR-2002; 2002WO-US009671.  
PF  
XX  
XX 28-MAR-2001; 2001US-0279495P.  
PR  
XX  
XX 21-MAY-2001; 2001US-0292544P.  
PR  
XX  
XX 08-AUG-2001; 2001US-0310801P.  
PR  
XX  
XX 01-OCT-2001; 2001US-0326370P.  
PR  
XX  
XX 04-DEC-2001; 2001US-0336780P.  
PR  
XX  
XX 20-FEB-2002; 2002US-0358985P.  
PR  
XX  
XX (ZYCO-) ZYCO INC.  
PA  
XX  
XX Chicz RM, Tomlinson AJ, Urban RG;  
PI  
XX  
XX WPI, 2003-040607/03.  
DR

XX New polypeptides (e.g. kinases, phosphatases, proteases, transporters,  
PT cytoskeletal proteins, receptors or transcription factors), useful for  
PT treating cancer, e.g. colon cancer, gastric cancer, sarcoma, lymphoma or  
PT leukemia.

XX Example 2; SEQ ID NO 1873; 134pp; English.

CC The invention describes a purified polypeptide, which comprises a  
CC fragment of a kinase, phosphatase, protease, protease inhibitor,  
CC transporter, cytoskeletal protein, receptor or transcription factor. The  
CC polypeptide is useful as an immunogenic composition for eliciting in a  
CC mammal an immunogenic response directed against any of the purified  
CC polypeptide. The purified polypeptide, or the antibody that binds to this  
CC polypeptide, is useful for treating cancer. The polypeptide is also  
CC useful for identifying compounds that binds to a naturally processed  
CC class I or class II MHC-binding polypeptide. The polypeptides and  
CC polynucleotides are particularly useful for treating or preventing  
CC myeloma, colon cancer, gastric cancer, adenocarcinoma, sarcoma, melanoma,  
CC lymphoma or leukemia. These are also useful for screening agents for  
CC treating the above mentioned diseases. This sequence represents an  
CC expressed protein tag (EPT) isolated from human tissue for translational  
CC profiling. Note: This sequence does not appear in the printed  
CC specification but was obtained in electronic format directly from WIPO at  
CC ftp.wipo.int/pub/published\_pct\_sequences

XX Sequence 524 AA;

Query Match 100.0%; Score 2789; DB 6; Length 524;  
Best Local Similarity 100.0%; Pred. No. 2,3e-237;  
Matches 524; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MYALFLASLLGALAGVILGLKECTRGSAVWCNVKTASDCGAVKHCLQTVWKNKPTVKS 60  
DB 1 MYALFLASLLGALAGVILGLKECTRGSAVWCNVKTASDCGAVKHCLQTVWKNKPTVKS 60  
QY 61 LPDCICKDVTAAGDMLKDNATEEELIYLEKTCQMLPKPMNSASCKEIVSYLPVILI 120  
DB 61 LPDCICKDVTAAGDMLKDNATEEELIYLEKTCQMLPKPMNSASCKEIVSYLPVILI 120  
QY 121 IKGEMSRGEVCSALNLCESLQKHIAELNHOKLESNKIPELDMTEVVAPEMANIPLLLY 180  
DB 121 IKGEMSRGEVCSALNLCESLQKHIAELNHOKLESNKIPELDMTEVVAPEMANIPLLLY 180

Db	121	IKGEMSRGCEVCSALNICESTIQKHLAEHLNHOXOLESNKIPELDMTEVAPFEMANIPLILY	180
Qy	181	PODGRSKXPBKNDGVOCODCIQWVTDIQTAVRNTSTFVQALVEHVEECRLRPGNADI	240
Db	181	PODGRSKXPCKNDGVOCODCIQWVTDIQTAVRNTSTFVQALVEHVEECRLRPGNADI	240
Qy	241	CKNTISQVSEFALIQMMHMQPKETCALVGFCDVEVEMPOQTLVPAKVASKNVIPALELVE	300
Db	241	CKNTISQVSEFALIQMMHMQPKETCALVGFCDVEVEMPOQTLVPAKVASKNVIPALELVE	300
Qy	301	PIKHEVPAKSDVYCEVCEFLVKEVTKLIDNNKTEKELIDADFKKCSLPSKJSEECOEY	360
Db	301	PIKHEVPAKSDVYCEVCEFLVKEVTKLIDNNKTEKELIDADFKKCSLPSKJSEECOEY	360
Qy	361	VDIYGGSSILSLLEEVSPBLVCSMHLCSGTRLPALTHAVTQPKDGGCEVCKKLYGYLD	420
Db	361	VDIYGGSSILSLLEEVSPBLVCSMHLCSGTRLPALTHAVTQPKDGGCEVCKKLYGYLD	420
Qy	421	RNEKSTKXELIILAIEKGCSPLPDYPQKQCFVAEYEPVLIEIIVNDPSPVCLKIG	480
Db	421	RNEKSTKXELIILAIEKGCSPLPDYPQKQCFVAEYEPVLIEIIVNDPSPVCLKIG	480
Qy	481	ACPSAHKPELIGTEKCIWGPSYWCNTRTAQOCNAVHEKRRVWN	524
Db	481	ACPSAHKPELIGTEKCIWGPSYWCNTRTAQOCNAVHEKRRVWN	524

transposase, cytoskeletal protein, receptor or transcription factor. The polypeptide is useful as an immunogenic composition for eliciting an immune response directed against any of the purified polypeptide. The purified polypeptide, or the antibody that binds to this polypeptide, is useful for treating cancer. The polypeptide is also useful for identifying compounds that binds to a naturally processed class I or class II MHC-binding polypeptide. The polypeptides and polynucleotides are particularly useful for treating or preventing myeloma, colon cancer, gastric cancer, adenocarcinoma, sarcoma, melanoma, lymphoma or leukaemia. These are also useful for screening agents for treating the above mentioned diseases. This sequence represents an expressed protein tag (EPT) isolated from human tissue for translational profiling. Note: This sequence does not appear in the printed specification but was obtained in electronic format directly from WIPO at [ftp://wipo.int/pub/published\\_pct\\_sequences](http://wipo.int/pub/published_pct_sequences)

KW adenocarcinoma; sarcoma; melanoma; lymphoma; leukaemia.  
 XX Homo sapiens.  
 OS  
 XX MO200278524-A2.  
 PN  
 XX 10-OCT-2002.  
 PD  
 XX 28-MAR-2002; 2002WO-US009671.  
 PF  
 XX 28-MAR-2001; 2001US-0279495P.  
 PR 21-MAY-2001; 2001US-0292544P.  
 PR 08-AUG-2001; 2001US-0310801P.  
 PR 01-OCT-2001; 2001US-0326370P.  
 PR 04-DEC-2001; 2001US-0336780P.  
 PR 20-FEB-2002; 2002US-0358985P.  
 XX  
 XX (ZYCO-) ZYCOS INC.  
 PA  
 XX Chiciz RM, Tomlinson AJ, Urban RG;  
 PI  
 XX WPI; 2003-040607/03.  
 DR  
 XX New polypeptides (e.g. kinases, phosphatases, proteases, transporters,  
 PT cytoskeletal proteins, receptors or transcription factors), useful for  
 PT treating cancer, e.g. colon cancer, gastric cancer, sarcoma, lymphoma or  
 PT leukemia.  
 XX  
 XX Example 2; SEQ ID NO 2041; 134pp; English.  
 PS  
 XX The invention describes a purified polypeptide, which comprises a  
 CC fragment of a kinase, phosphatase, protease, protease inhibitor,  
 CC transporter, cytoskeletal protein, receptor or transcription factor. The  
 CC polypeptide is useful as an immunogenic composition for eliciting in a  
 CC mammal an immunogenic response directed against any of the purified  
 CC polypeptide. The purified polypeptide, or the antibody that binds to this  
 CC polypeptide, is useful for treating cancer. The polypeptide is also  
 CC useful for identifying compounds that binds to a naturally processed  
 CC class I or class II MHC-binding polypeptide. The polypeptides and  
 CC polynucleotides are particularly useful for treating or preventing  
 CC myeloma, colon cancer, gastric cancer, adenocarcinoma, sarcoma, melanoma,  
 CC lymphoma or leukaemia. These are also useful for screening agents for  
 CC treating the above mentioned diseases. This sequence represents an  
 CC expressed protein tag (EPT) isolated from human tissue for translational  
 CC profiling. Note: This sequence does not appear in the printed  
 CC specification but was obtained in electronic format directly from WIPO at  
 CC ftp.wipo.int/pub/published\_pct\_sequences  
 CC  
 XX Sequence 524 AA:  
 SQ  
 Query Match 100.0%; Score 2789; DB 6; Length 524;  
 Best Local Similarity 100.0%; Pred. No. 2.3e-237;  
 Matches 524; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 301 PIKHEVPAKSDVYCEVCEFLVKEVTKLIDNNKTEKEILDAFDKWCSTLPSLSSECOEV 360  
 DB 301 PIKHEVPAKSDVYCEVCEFLVKEVTKLIDNNKTEKEILDAFDKWCSTLPSLSSECOEV 360  
 QY 361 VDTYGSILSLIEEVPPELVCSMLHCSGTRLPALTVHTVTPQPDGFCVEYCKLVGYLD 420  
 DB 361 VDTYGSILSLIEEVPPELVCSMLHCSGTRLPALTVHTVTPQPDGFCVEYCKLVGYLD 420  
 QY 421 RNLEKSTKOEIILAELKSGSFLPDYPYQKQDGFVAEYEPVLIILVWDPSTVCLKIG 480  
 DB 421 RNLEKSTKOEIILAELKSGSFLPDYPYQKQDGFVAEYEPVLIILVWDPSTVCLKIG 480  
 QY 481 ACPSAHKPLGTGECICWSPSYWCNTETAACNAVEHCKRHVN 524  
 DB 481 ACPSAHKPLGTGECICWSPSYWCNTETAACNAVEHCKRHVN 524  
 RESULT 9  
 AB05216  
 ID AB05216 standard; protein; 524 AA.  
 XX AC  
 XX AB05216;  
 XX  
 XX 29-JAN-2003 (first entry)  
 DT  
 XX Human expressed protein tag (EPT) #1882.  
 DE  
 XX Translational profiling; expressed protein tag; EPT; kinase; phosphatase;  
 KW protease; protease inhibitor; transporter; cytoskeletal protein;  
 KW receptor; transcription factor; cancer; MHC;  
 KW major histocompatibility complex; myeloma; colon cancer; gastric cancer;  
 KW adenocarcinoma; sarcoma; melanoma; lymphoma; leukaemia.  
 XX  
 XX Homo sapiens.  
 OS  
 XX MO200278524-A2.  
 PN  
 XX 10-OCT-2002.  
 PD  
 XX 28-MAR-2002; 2002WO-US009671.  
 PF  
 XX 28-MAR-2001; 2001US-0279495P.  
 PR 21-MAY-2001; 2001US-0292544P.  
 PR 08-AUG-2001; 2001US-0310801P.  
 PR 01-OCT-2001; 2001US-0326370P.  
 PR 04-DEC-2001; 2001US-0336780P.  
 PR 20-FEB-2002; 2002US-0358985P.  
 XX  
 XX (ZYCO-) ZYCOS INC.  
 PA  
 XX Chiciz RM, Tomlinson AJ, Urban RG;  
 PI  
 XX WPI; 2003-040607/03.  
 DR  
 XX New polypeptides (e.g. kinases, phosphatases, proteases, transporters,  
 PT cytoskeletal proteins, receptors or transcription factors), useful for  
 PT treating cancer, e.g. colon cancer, gastric cancer, sarcoma, lymphoma or  
 PT leukemia.  
 XX  
 XX Example 2; SEQ ID NO 1882; 134pp; English.  
 PS  
 XX The invention describes a purified polypeptide, which comprises a  
 CC fragment of a kinase, phosphatase, protease, protease inhibitor,  
 CC transporter, cytoskeletal protein, receptor or transcription factor. The  
 CC polypeptide is useful as an immunogenic composition for eliciting in a  
 CC mammal an immunogenic response directed against any of the purified  
 CC polypeptide. The purified polypeptide, or the antibody that binds to this  
 CC polypeptide, is useful for treating cancer. The polypeptide is also  
 CC useful for identifying compounds that binds to a naturally processed  
 CC class I or class II MHC-binding polypeptide. The polypeptides and  
 CC polynucleotides are particularly useful for treating or preventing  
 CC myeloma, colon cancer, gastric cancer, adenocarcinoma, sarcoma, melanoma,

CC lymphoma or leukaemia. These are also useful for screening agents for  
 CC treating the above mentioned diseases. This sequence represents an  
 CC expressed protein tag (EPT) isolated from human tissue for translational  
 CC profiling. Note: this sequence does not appear in the printed  
 CC specification but was obtained in electronic format directly from WIPO at  
 CC ftp.wipo.int/pub/published\_pct\_sequences

XX Sequence 524 AA;

Query Match 100.0%; Score 2789; DB 6; Length 524;  
 Best Local Similarity 100.0%; Pred. No. 2.3e-237;  
 Matches 524; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY 1 MYALFLIASLIGALAGPVGLGKECTRGSAVWCQNVKTAASDCGAVKCLQTVMNKPVTYS 60
DB 1 MYALFLIASLIGALAGPVGLGKECTRGSAVWCQNVKTAASDCGAVKCLQTVMNKPVTYS 60
QY 61 LPEDICKDVVTAAGDMKDNATEEELIYVLEKTCMWLPKPNMSASCKEIVDSYLPVILDI 120
DB 61 LPEDICKDVVTAAGDMKDNATEEELIYVLEKTCMWLPKPNMSASCKEIVDSYLPVILDI 120
QY 121 IKGEMSRPGEVCSALNLCESLQKHLAELNHQKLESNKIPELDMEVAPFMANIPILLY 180
DB 121 IKGEMSRPGEVCSALNLCESLQKHLAELNHQKLESNKIPELDMEVAPFMANIPILLY 180
QY 181 PODGPRSKPQPKDNGDVQDCIQWVTDIQTAVRINSTFVQALVHVKECDRLGPGMADI 240
DB 181 PODGPRSKPQPKDNGDVQDCIQWVTDIQTAVRINSTFVQALVHVKECDRLGPGMADI 240
QY 241 CKNYISQYSEIAIQMMHMQPEKICALVGFCDVEKEMPMQTLVPAKVASKNVITALEVE 300
DB 241 CKNYISQYSEIAIQMMHMQPEKICALVGFCDVEKEMPMQTLVPAKVASKNVITALEVE 300
QY 301 PIKKEHVPAKSDVYCEVCEFLVKEVTKLIDNNKTEKEIIDAFDKMCSKLPKLSSECOEV 360
DB 301 PIKKEHVPAKSDVYCEVCEFLVKEVTKLIDNNKTEKEIIDAFDKMCSKLPKLSSECOEV 360
QY 361 VDTYSSILSLILEEVSPELVCSMLHLCSGTRLPALTVHVTQPKDGFCEVCKLVGYLD 420
DB 361 VDTYSSILSLILEEVSPELVCSMLHLCSGTRLPALTVHVTQPKDGFCEVCKLVGYLD 420
QY 421 RNLEKNSTKOEIILALEKGCSTFLPDPYOKCQDQFAVEYEPVILIVEVMDPSFVCLKIG 480
DB 421 RNLEKNSTKOEIILALEKGCSTFLPDPYOKCQDQFAVEYEPVILIVEVMDPSFVCLKIG 480
QY 481 ACPSSAHKPLGTGECTWGPSTWCQNTETAAQCNAVEHCKRHVMN 524
DB 481 ACPSSAHKPLGTGECTWGPSTWCQNTETAAQCNAVEHCKRHVMN 524

```

RESULT 10  
 ID AB005202 standard; protein, 524 AA.

XX AB005202;  
 XX 29-JAN-2003 (first entry)

DE Human expressed protein tag (EPT) #1868.

XX Translational profiling; expressed protein tag; EPT; kinase; phosphatase;  
 KM protease; protease inhibitor; transporter; cytoskeletal protein;  
 KM receptor; transcription factor; cancer; MHC;  
 KM major histocompatibility complex; myeloma; colon cancer; gastric cancer;  
 KM adenocarcinoma; sarcoma; melanoma; lymphoma; leukaemia.

OS Homo sapiens.  
 XX PN WO200278524-A2.  
 XX 10-OCT-2002.  
 XX PD 28-MAR-2002; 2002WO-US009671.  
 XX PF

XX 28-MAR-2001; 2001US-0279495P.  
 PR 21-MAY-2001; 2001US-0292544P.  
 PR 08-AUG-2001; 2001US-0310801P.  
 PR 01-OCT-2001; 2001US-0326370P.  
 PR 04-DEC-2001; 2001US-0336780P.  
 PR 20-FEB-2002; 2002US-0358985P.

XX (ZYCO-) ZYCO INC.

XX Chicx RM, Tomlinson AJ, Urban RG;

XX WPI; 2003-040607/03.

PT New polypeptides (e.g. kinases, phosphatases, proteases, transporters,  
 PT cytoskeletal proteins, receptors or transcription factors), useful for  
 PT treating cancer, e.g. colon cancer, gastric cancer, sarcoma, lymphoma or  
 PT leukaemia.

XX Example 2; SEQ ID NO 1868; 134bp; English.

CC The invention describes a purified polypeptide, which comprises a  
 CC fragment of a kinase, phosphatase, protease, protease inhibitor,  
 CC transporter, cytoskeletal protein, receptor or transcription factor. The  
 CC polypeptide is useful as an immunogenic composition for eliciting in a  
 CC mammal an immunogenic response directed against any of the purified  
 CC polypeptide, is useful for treating cancer. The polypeptide is also  
 CC useful for identifying compounds that binds to a naturally processed  
 CC class I or class II MHC-binding polypeptide. The polypeptides and  
 CC polynucleotides are particularly useful for treating or preventing  
 CC myeloma, colon cancer, gastric cancer, adenocarcinoma, sarcoma, melanoma,  
 CC lymphoma or leukaemia. These are also useful for screening agents for  
 CC treating the above mentioned diseases. This sequence represents an  
 CC expressed protein tag (EPT) isolated from human tissue for translational  
 CC profiling. Note: This sequence does not appear in the printed  
 CC specification but was obtained in electronic format directly from WIPO at  
 CC ftp.wipo.int/pub/published\_pct\_sequences

XX Sequence 524 AA;

Query Match 100.0%; Score 2789; DB 6; Length 524;  
 Best Local Similarity 100.0%; Pred. No. 2.3e-237;  
 Matches 524; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY 1 MYALFLIASLIGALAGPVGLGKECTRGSAVWCQNVKTAASDCGAVKCLQTVMNKPVTYS 60
DB 1 MYALFLIASLIGALAGPVGLGKECTRGSAVWCQNVKTAASDCGAVKCLQTVMNKPVTYS 60
QY 61 LPEDICKDVVTAAGDMKDNATEEELIYVLEKTCMWLPKPNMSASCKEIVDSYLPVILDI 120
DB 61 LPEDICKDVVTAAGDMKDNATEEELIYVLEKTCMWLPKPNMSASCKEIVDSYLPVILDI 120
QY 121 IKGEMSRPGEVCSALNLCESLQKHLAELNHQKLESNKIPELDMEVAPFMANIPILLY 180
DB 121 IKGEMSRPGEVCSALNLCESLQKHLAELNHQKLESNKIPELDMEVAPFMANIPILLY 180
QY 181 PODGPRSKPQPKDNGDVQDCIQWVTDIQTAVRINSTFVQALVHVKECDRLGPGMADI 240
DB 181 PODGPRSKPQPKDNGDVQDCIQWVTDIQTAVRINSTFVQALVHVKECDRLGPGMADI 240
QY 241 CKNYISQYSEIAIQMMHMQPEKICALVGFCDVEKEMPMQTLVPAKVASKNVITALEVE 300
DB 241 CKNYISQYSEIAIQMMHMQPEKICALVGFCDVEKEMPMQTLVPAKVASKNVITALEVE 300
QY 301 PIKKEHVPAKSDVYCEVCEFLVKEVTKLIDNNKTEKEIIDAFDKMCSKLPKLSSECOEV 360
DB 301 PIKKEHVPAKSDVYCEVCEFLVKEVTKLIDNNKTEKEIIDAFDKMCSKLPKLSSECOEV 360
QY 361 VDTYSSILSLILEEVSPELVCSMLHLCSGTRLPALTVHVTQPKDGFCEVCKLVGYLD 420
DB 361 VDTYSSILSLILEEVSPELVCSMLHLCSGTRLPALTVHVTQPKDGFCEVCKLVGYLD 420

```

QY 421 RNLEKNSKTQKQIILALEKGCSEFLPDPYQKQCDQFVAEYEPVLIILVEMDPSFVCLKIG 480  
Db 421 RNLEKNSKTQKQIILALEKGCSEFLPDPYQKQCDQFVAEYEPVLIILVEMDPSFVCLKIG 480  
QY 481 ACPSAHKRPLLGTEKCIWGPSYWCNTETPAACNAVEHCKRHYVN 524  
Db 481 ACPSAHKRPLLGTEKCIWGPSYWCNTETPAACNAVEHCKRHYVN 524

RESULT 11  
ABU05214  
ID ABU05214 standard; protein; 524 AA.  
XX  
AC ABU05214;  
XX  
DT 29-JAN-2003 (first entry)  
XX  
DE Human expressed protein tag (EPT) #1880.  
XX  
KW Translational profiling; expressed protein tag; EPT; kinase; phosphatase;  
KW protease; protease inhibitor; transporter; cytoskeletal protein;  
KW receptor; transcription factor; cancer; MHC;  
KW major histocompatibility complex; myeloma; colon cancer; gastric cancer;  
KW adenocarcinoma; sarcoma; melanoma; lymphoma; leukaemia.  
XX  
OS Homo sapiens.  
XX  
FN WO200278524-A2.  
XX  
PD 10-OCT-2002.  
XX  
PE 28-MAR-2002; 2002WO-US009671.  
XX  
PR 28-MAR-2001; 2001US-0279495P.  
PR 21-MAY-2001; 2001US-0292544P.  
PR 08-AUG-2001; 2001US-0310801P.  
PR 01-OCT-2001; 2001US-0326370P.  
PR 04-DEC-2001; 2001US-0336780P.  
PR 20-FEB-2002; 2002US-0358985P.  
XX  
PA (ZYCO-) ZYCOS INC.  
PI Chicx RM, Tomlinson AJ, Urban RG;  
XX  
DR WPI; 2003-040607/03.  
XX  
PT New polypeptides (e.g. kinases, phosphatases, proteases, transporters,  
PT cytoskeletal proteins, receptors or transcription factors), useful for  
PT treating cancer, e.g. colon cancer, gastric cancer, sarcoma, lymphoma or  
PT leukemia.  
XX  
PS Example 2; SEQ ID NO 1880; 134pp; English.  
XX  
CC The invention describes a purified polypeptide, which comprises a  
CC fragment of a kinase, phosphatase, protease, protease inhibitor,  
CC transporter, cytoskeletal protein, receptor or transcription factor. The  
CC polypeptide is useful as an immunogenic composition for eliciting in a  
CC mammal an immunogenic response directed against any of the purified  
CC polypeptide. The purified polypeptide, or the antibody that binds to this  
CC polypeptide, is useful for treating cancer. The polypeptide is also  
CC useful for identifying compounds that binds to a naturally processed  
CC class I or class II MHC-binding polypeptide. The polypeptides and  
CC polynucleotides are particularly useful for treating or preventing  
CC myeloma, colon cancer, gastric cancer, adenocarcinoma, sarcoma, melanoma,  
CC lymphoma or leukaemia. These are also useful for screening agents for  
CC treating the above mentioned diseases. This sequence represents an  
CC expressed protein tag (EPT) isolated from human tissue for translational  
CC profiling. Note: This sequence does not appear in the printed  
CC specification but was obtained in electronic format directly from WIPO at  
CC ftp.wipo.int/pub/published\_pct\_sequences  
XX  
SQ Sequence 524 AA;

Query Match 100.0%; Score 2789; DB 6; Length 524;  
Best Local Similarity 100.0%; Pred. No. 2,3e-237;  
Matches 524; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MYALFLASLIGALAGVVLGLKECTRGSAVWCONVKTASDCGAVKICLQTVNKKPTVKS 60  
Db 1 MYALFLASLIGALAGVVLGLKECTRGSAVWCONVKTASDCGAVKICLQTVNKKPTVKS 60  
QY 61 LPDICI CKDVVTAAGDMLKDNATEEELVYLEKTCDDMLPKPMNSASCKEIVDSYLPVILI 120  
Db 61 LPDICI CKDVVTAAGDMLKDNATEEELVYLEKTCDDMLPKPMNSASCKEIVDSYLPVILI 120  
QY 121 IKGMSRPGEVCSALNICESIQKHLAEINHOQKLESNKIPBLDMTEVVAAPMANIPILLY 180  
Db 121 IKGMSRPGEVCSALNICESIQKHLAEINHOQKLESNKIPBLDMTEVVAAPMANIPILLY 180  
QY 181 PDGPRSKPQKNDGVCQDCIQWVTDIQTAVRTNSTFVQALVBEHVECDRLGPGMADI 240  
Db 181 PDGPRSKPQKNDGVCQDCIQWVTDIQTAVRTNSTFVQALVBEHVECDRLGPGMADI 240  
QY 241 CKNTISQYSEIATQMMHMQPEICALVGFCDVEKEMQTLVPAKVASKVIVIPALIVE 300  
Db 241 CKNTISQYSEIATQMMHMQPEICALVGFCDVEKEMQTLVPAKVASKVIVIPALIVE 300  
QY 301 PIKKHEVPAKSDVYCEVEFLVKEVTYKIDNNKTEKEILDAFDKVCSTLPSLSBECQEV 360  
Db 301 PIKKHEVPAKSDVYCEVEFLVKEVTYKIDNNKTEKEILDAFDKVCSTLPSLSBECQEV 360  
QY 361 VDTYGSSTLSLLEBSEVELVCSMLHCSGTRLPALTAVHVPQKDGCEVCKLVGLD 420  
Db 361 VDTYGSSTLSLLEBSEVELVCSMLHCSGTRLPALTAVHVPQKDGCEVCKLVGLD 420  
QY 421 RNLEKNSKTQKQIILALEKGCSEFLPDPYQKQCDQFVAEYEPVLIILVEMDPSFVCLKIG 480  
Db 421 RNLEKNSKTQKQIILALEKGCSEFLPDPYQKQCDQFVAEYEPVLIILVEMDPSFVCLKIG 480  
QY 481 ACPSAHKRPLLGTEKCIWGPSYWCNTETPAACNAVEHCKRHYVN 524  
Db 481 ACPSAHKRPLLGTEKCIWGPSYWCNTETPAACNAVEHCKRHYVN 524

RESULT 12  
ABU05215  
ID ABU05215 standard; protein; 524 AA.  
XX  
AC ABU05215;  
XX  
DT 29-JAN-2003 (first entry)  
XX  
DE Human expressed protein tag (EPT) #1881.  
XX  
KW Translational profiling; expressed protein tag; EPT; kinase; phosphatase;  
KW protease; protease inhibitor; transporter; cytoskeletal protein;  
KW receptor; transcription factor; cancer; MHC;  
KW major histocompatibility complex; myeloma; colon cancer; gastric cancer;  
KW adenocarcinoma; sarcoma; melanoma; lymphoma; leukaemia.  
XX  
OS Homo sapiens.  
XX  
FN WO200278524-A2.  
XX  
PD 10-OCT-2002.  
XX  
PE 28-MAR-2002; 2002WO-US009671.  
XX  
PR 28-MAR-2001; 2001US-0279495P.  
PR 21-MAY-2001; 2001US-0292544P.  
PR 08-AUG-2001; 2001US-0310801P.  
PR 01-OCT-2001; 2001US-0326370P.  
PR 04-DEC-2001; 2001US-0336780P.  
PR 20-FEB-2002; 2002US-0358985P.  
XX  
PA (ZYCO-) ZYCOS INC.



XX Chicz RM, Tomlinson AJ, Urban RG;  
 XX WPI; 2003-040607/03.  
 PT New polypeptides (e.g. kinases, phosphatases, proteases, transporters,  
 PT cytoskeletal proteins, receptors or transcription factors), useful for  
 PT treating cancer, e.g. colon cancer, gastric cancer, sarcoma, lymphoma or  
 XX leukemia.  
 PS Example 2; SEQ ID NO 1881; 134pp; English.  
 CC The invention describes a purified polypeptide, which comprises a  
 CC fragment of a kinase, phosphatase, protease, protease inhibitor,  
 CC transporter, cytoskeletal protein, receptor or transcription factor. The  
 CC polypeptide is useful as an immunogenic composition for eliciting in a  
 CC mammal an immunogenic response directed against any of the purified  
 CC polypeptide, the purified polypeptide, or the antibody that binds to this  
 CC useful for identifying compounds that binds to a naturally processed  
 CC class I or class II MHC-binding polypeptide. The polypeptide is also  
 CC polynucleotides are particularly useful for treating or preventing  
 CC myeloma, colon cancer, gastric cancer, adenocarcinoma, sarcoma, melanoma,  
 CC lymphoma or leukemia. These are also useful for screening agents for  
 CC treating the above mentioned diseases. This sequence represents an  
 CC expressed protein tag (EPT) isolated from human tissue for translational  
 CC profiling. Note: This sequence does not appear in the printed  
 CC specification but was obtained in electronic format directly from WIPO at  
 CC ftp.wipo.int/pub/published\_pct\_sequences  
 XX Sequence 524 AA;

Query Match 100.0%; Score 2789; DB 6; Length 524;  
 Best Local Similarity 100.0%; Pred. No. 2.3e-237;  
 Matches 524; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY 1 MYALFLASLIGALAGVLTGLKECTRGSAVWCQNVKTASDCGAVKHCLOTVMNKPPTVKS 60
DB 1 MYALFLASLIGALAGVLTGLKECTRGSAVWCQNVKTASDCGAVKHCLOTVMNKPPTVKS 60
QY 61 LPDCICKDVTTAAGDMLKDNATEEIIIVYLEKTCDMLPKPNMSASCKEIVSYLPVILI 120
DB 61 LPDCICKDVTTAAGDMLKDNATEEIIIVYLEKTCDMLPKPNMSASCKEIVSYLPVILI 120
QY 121 IKGEMSRGGEVCSALNLCESLQKHLAEINHQKOLESNKIPELDMTEVAVPMMANIPILLY 180
DB 121 IKGEMSRGGEVCSALNLCESLQKHLAEINHQKOLESNKIPELDMTEVAVPMMANIPILLY 180
QY 181 FODGPRSRKPPKNDGVQDCICQWVTDIQTAVRNTSTFVQALVHVKEECRLGPGMADI 240
DB 181 FODGPRSRKPPKNDGVQDCICQWVTDIQTAVRNTSTFVQALVHVKEECRLGPGMADI 240
QY 241 CKNTIISQYSEIATIGMMHMPKEICALVGFCEVKEPMQTLVPAKASKNVIPALERVE 300
DB 241 CKNTIISQYSEIATIGMMHMPKEICALVGFCEVKEPMQTLVPAKASKNVIPALERVE 300
QY 301 PIKGEVPAKSDVYCEVCEFLVKEVTKLIDNNKTEKILIDAFDMCKSKLPSISEEQEV 360
DB 301 PIKGEVPAKSDVYCEVCEFLVKEVTKLIDNNKTEKILIDAFDMCKSKLPSISEEQEV 360
QY 361 VDIYSSILSLLEEVPELVCSMLHCSGTRLPALVHTVTPDDGFCVCCKLVGYLD 420
DB 361 VDIYSSILSLLEEVPELVCSMLHCSGTRLPALVHTVTPDDGFCVCCKLVGYLD 420
QY 421 RNFENKSTKQELLALAEKGCFLPDPYQKODQFVAEYEPVLIELIVEMDPSFVCLKTG 480
DB 421 RNFENKSTKQELLALAEKGCFLPDPYQKODQFVAEYEPVLIELIVEMDPSFVCLKTG 480
QY 481 ACSANRKLGTETECIMGPSYWCQNTETAACNAVHECKRHVNN 524
DB 481 ACSANRKLGTETECIMGPSYWCQNTETAACNAVHECKRHVNN 524

```

RESULT 13  
 ABU05199  
 ID ABU05199 standard; protein; 524 AA.  
 XX  
 AC ABU05199;  
 XX  
 DT 29-JAN-2003 (first entry)  
 XX  
 DE Human expressed protein tag (EPT) #1865.  
 XX  
 KW Translational profiling; expressed protein tag; EPT; kinase; phosphatase;  
 KW protease; protease inhibitor; transporter; cytoskeletal protein;  
 KW receptor; transcription factor; cancer; MHC;  
 KW major histocompatibility complex; myeloma; colon cancer; gastric cancer;  
 KW adenocarcinoma; sarcoma; melanoma; lymphoma; leukemia.  
 OS Homo sapiens.  
 XX  
 PN WO200278524-A2.  
 XX  
 PD 10-OCT-2002.  
 XX  
 PF 28-MAR-2002; 2002WO-US009671.  
 XX  
 PR 28-MAR-2001; 2001US-0279495P.  
 PR 21-MAY-2001; 2001US-0292544P.  
 PR 08-AUG-2001; 2001US-0310801P.  
 PR 01-OCT-2001; 2001US-0326370P.  
 PR 04-DEC-2001; 2001US-0336780P.  
 PR 20-FEB-2002; 2002US-0358985P.  
 XX  
 PA (ZYCO-) ZYCOS INC.  
 XX  
 PT Chicz RM, Tomlinson AJ, Urban RG;  
 PT WPI; 2003-040607/03.  
 DR  
 XX  
 PT New polypeptides (e.g. kinases, phosphatases, proteases, transporters,  
 PT cytoskeletal proteins, receptors or transcription factors), useful for  
 PT treating cancer, e.g. colon cancer, gastric cancer, sarcoma, lymphoma or  
 PT leukemia.  
 XX  
 PS Example 2; SEQ ID NO 1865; 134pp; English.  
 CC The invention describes a purified polypeptide, which comprises a  
 CC fragment of a kinase, phosphatase, protease, protease inhibitor,  
 CC transporter, cytoskeletal protein, receptor or transcription factor. The  
 CC polypeptide is useful as an immunogenic composition for eliciting in a  
 CC mammal an immunogenic response directed against any of the purified  
 CC polypeptide, the purified polypeptide, or the antibody that binds to this  
 CC useful for identifying compounds that binds to a naturally processed  
 CC class I or class II MHC-binding polypeptide. The polypeptide is also  
 CC polynucleotides are particularly useful for treating or preventing  
 CC myeloma, colon cancer, gastric cancer, adenocarcinoma, sarcoma, melanoma,  
 CC lymphoma or leukemia. These are also useful for screening agents for  
 CC treating the above mentioned diseases. This sequence represents an  
 CC expressed protein tag (EPT) isolated from human tissue for translational  
 CC profiling. Note: This sequence does not appear in the printed  
 CC specification but was obtained in electronic format directly from WIPO at  
 CC ftp.wipo.int/pub/published\_pct\_sequences  
 XX Sequence 524 AA;

Query Match 100.0%; Score 2789; DB 6; Length 524;  
 Best Local Similarity 100.0%; Pred. No. 2.3e-237;  
 Matches 524; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY 1 MYALFLASLIGALAGVLTGLKECTRGSAVWCQNVKTASDCGAVKHCLOTVMNKPPTVKS 60
DB 1 MYALFLASLIGALAGVLTGLKECTRGSAVWCQNVKTASDCGAVKHCLOTVMNKPPTVKS 60
QY 61 LPDCICKDVTTAAGDMLKDNATEEIIIVYLEKTCDMLPKPNMSASCKEIVSYLPVILI 120

```



```
Db      61  LPCDICKDVTTAAGDMLKDNATEEELVYLEKTCOMLEKPMMSCKEIVSYLPIVLI 120
Qy      121  IKGEMSRGGEVCSALNLCESIQKHIAELNHOKOLESNKIPELDMTEVAPPMANIPILLY 180
Db      121  IKGEMSRGGEVCSALNLCESIQKHIAELNHOKOLESNKIPELDMTEVAPPMANIPILLY 180
Qy      181  PDDGPRSKPQPKDNGDVQCDCIQWVTDIQTAVRTNSTFVQALVHEVKEECRLGGMADI 240
Db      181  PDDGPRSKPQPKDNGDVQCDCIQWVTDIQTAVRTNSTFVQALVHEVKEECRLGGMADI 240
Qy      241  CKNTISQYSEIAIQMMHMOPEKICALVGFCDVEKEMQTLVPAKASKVIVIPALIEVE 300
Db      241  CKNTISQYSEIAIQMMHMOPEKICALVGFCDVEKEMQTLVPAKASKVIVIPALIEVE 300
Qy      301  PIKHEVPKSDVYCEVCEFLVKEVTKLIDNNKTEKEILDADFDMKCSKLPKSLSECOEV 360
Db      301  PIKHEVPKSDVYCEVCEFLVKEVTKLIDNNKTEKEILDADFDMKCSKLPKSLSECOEV 360
Qy      361  VDTYGSSTLILBEVSPELVCSMLHLCSGTRLPALTVAHTVQPKDGGFCVCKLVGYLD 420
Db      361  VDTYGSSTLILBEVSPELVCSMLHLCSGTRLPALTVAHTVQPKDGGFCVCKLVGYLD 420
Qy      421  RNLEKNSTKOEIILALEKGCSEFLPDPYOKCDOQFAVEYEPVILIEIVMNDPSFYCLKIG 480
Db      421  RNLEKNSTKOEIILALEKGCSEFLPDPYOKCDOQFAVEYEPVILIEIVMNDPSFYCLKIG 480
Qy      481  ACPSAHKEPLGTERKICMGPSYWCNTETFAACNAVEHCKRHVMN 524
Db      481  ACPSAHKEPLGTERKICMGPSYWCNTETFAACNAVEHCKRHVMN 524
```

## RESULT 14

ABU05212 ID ABU05212 standard; protein: 524 AA.

AC ABU05212;

DT 29-JAN-2003 (first entry)

XX Human expressed protein tag (EPT) #1878.

XX Translational profiling: expressed protein tag; EPT; kinase; phosphatase;

XX protease; protease inhibitor; transporter; cytoskeletal protein;

XX receptor; transcription factor; cancer; MHC; myeloma; colon cancer; gastric cancer;

XX adenocarcinoma; sarcoma; melanoma; lymphoma; leukaemia.

XX Homo sapiens.

XX WO200278524-A2.

XX 10-OCT-2002.

XX 28-MAR-2002; 2002WO-US009671.

XX 28-MAR-2001; 2001US-0279495P.

XX 21-MAY-2001; 2001US-0292544P.

XX 08-AUG-2001; 2001US-0310801P.

XX 01-OCT-2001; 2001US-0326370P.

XX 04-DEC-2001; 2001US-0336780P.

XX 20-FEB-2002; 2002US-0358985P.

XX (ZYCO-) ZYCOs INC.

XX Chiciz RM, Tomlinson AJ, Urban RG;

XX WPI; 2003-040607/03.

XX New polypeptides (e.g. kinases, phosphatases, proteases, transporters,

XX cytoskeletal proteins, receptors or transcription factors), useful for

XX treating cancer, e.g. colon cancer, gastric cancer, sarcoma, lymphoma or

XX leukemia.

```
XX      XX      Example 2; SEQ ID NO 1878; 134pp; English.
PS      CC      The invention describes a purified polypeptide, which comprises a
XX      CC      fragment of a kinase, phosphatase, protease, protease inhibitor,
CC      CC      transporter, cytoskeletal protein, receptor or transcription factor. The
CC      CC      polypeptide is useful as an immunogenic composition for eliciting in a
CC      CC      mammal an immunogenic response directed against any of the purified
CC      CC      polypeptide. The purified polypeptide, or the antibody that binds to this
CC      CC      polypeptide, is useful for treating cancer. The polypeptide is also
CC      CC      useful for identifying compounds that binds to a naturally processed
CC      CC      class I or class II MHC-binding polypeptide. The polypeptides and
CC      CC      polynucleotides are particularly useful for treating or preventing
CC      CC      myeloma, colon cancer, gastric cancer, adenocarcinoma, sarcoma, melanoma,
CC      CC      lymphoma or leukaemia. These are also useful for screening agents for
CC      CC      treating the above mentioned diseases. This sequence represents an
CC      CC      expressed protein tag (EPT) isolated from human tissue for translational
CC      CC      profiling. Note: This sequence does not appear in the printed
CC      CC      specification but was obtained in electronic format directly from WIPO at
CC      CC      ftp.wipo.int/pub/published_pct_sequences
SQ      Sequence 524 AA;
```

Query Match 100.0%; Score 2789; DB 6; Length 524;  
Best Local Similarity 100.0%; Pred. No. 2,3e-237;  
Matches 524; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy      1  MYALFLASLIGALAGVILGKECTRGSAVWCONVKTASDCGAVKCLQTVWNNKPTVKS 60
Db      1  MYALFLASLIGALAGVILGKECTRGSAVWCONVKTASDCGAVKCLQTVWNNKPTVKS 60
Qy      61  LPCDICKDVTTAAGDMLKDNATEEELVYLEKTCOMLEKPMMSCKEIVSYLPIVLI 120
Db      61  LPCDICKDVTTAAGDMLKDNATEEELVYLEKTCOMLEKPMMSCKEIVSYLPIVLI 120
Qy      121  IKGEMSRGGEVCSALNLCESIQKHIAELNHOKOLESNKIPELDMTEVAPPMANIPILLY 180
Db      121  IKGEMSRGGEVCSALNLCESIQKHIAELNHOKOLESNKIPELDMTEVAPPMANIPILLY 180
Qy      181  PDDGPRSKPQPKDNGDVQCDCIQWVTDIQTAVRTNSTFVQALVHEVKEECRLGGMADI 240
Db      181  PDDGPRSKPQPKDNGDVQCDCIQWVTDIQTAVRTNSTFVQALVHEVKEECRLGGMADI 240
Qy      241  CKNTISQYSEIAIQMMHMOPEKICALVGFCDVEKEMQTLVPAKASKVIVIPALIEVE 300
Db      241  CKNTISQYSEIAIQMMHMOPEKICALVGFCDVEKEMQTLVPAKASKVIVIPALIEVE 300
Qy      301  PIKHEVPKSDVYCEVCEFLVKEVTKLIDNNKTEKEILDADFDMKCSKLPKSLSECOEV 360
Db      301  PIKHEVPKSDVYCEVCEFLVKEVTKLIDNNKTEKEILDADFDMKCSKLPKSLSECOEV 360
Qy      361  VDTYGSSTLILBEVSPELVCSMLHLCSGTRLPALTVAHTVQPKDGGFCVCKLVGYLD 420
Db      361  VDTYGSSTLILBEVSPELVCSMLHLCSGTRLPALTVAHTVQPKDGGFCVCKLVGYLD 420
Qy      421  RNLEKNSTKOEIILALEKGCSEFLPDPYOKCDOQFAVEYEPVILIEIVMNDPSFYCLKIG 480
Db      421  RNLEKNSTKOEIILALEKGCSEFLPDPYOKCDOQFAVEYEPVILIEIVMNDPSFYCLKIG 480
Qy      481  ACPSAHKEPLGTERKICMGPSYWCNTETFAACNAVEHCKRHVMN 524
Db      481  ACPSAHKEPLGTERKICMGPSYWCNTETFAACNAVEHCKRHVMN 524
```

## RESULT 15

ABU05213 ID ABU05213 standard; protein: 524 AA.

AC ABU05213;

DT 29-JAN-2003 (first entry)

XX Human expressed protein tag (EPT) #1879.

XX Translational profiling; expressed protein tag; EPT, kinase, phosphatase;  
 KM protease; protease inhibitor; transporter; cytoskeletal protein;  
 KM receptor; transcription factor; cancer; MHC;  
 KM major histocompatibility complex; myeloma; colon cancer; gastric cancer;  
 KM adenocarcinoma; sarcoma; melanoma; lymphoma; leukemia.  
 OS Homo sapiens.  
 XX MO200278524-A2.  
 XX 10-OCT-2002.  
 XX 28-MAR-2002; 2002WO-US009671.  
 XX 28-MAR-2001; 2001US-0292549P.  
 PR 21-MAY-2001; 2001US-0292549P.  
 PR 08-AUG-2001; 2001US-0310801P.  
 PR 01-OCT-2001; 2001US-0326370P.  
 PR 04-DEC-2001; 2001US-0336780P.  
 PR 20-FEB-2002; 2002US-0358985P.  
 XX (ZYCO-) ZYCO INC.  
 PA Chicx RM, Tomlinson AJ, Urban RG;  
 PI WPI; 2003-040607/03.  
 DR New polypeptides (e.g. kinases, phosphatases, proteases, transporters,  
 XX cytoskeletal proteins, receptors or transcription factors), useful for  
 PT treating cancer, e.g. colon cancer, gastric cancer, sarcoma, lymphoma or  
 PT leukemia.  
 XX Example 2; SEQ ID NO 1879; 134pp; English.  
 XX The invention describes a purified polypeptide, which comprises a  
 CC fragment of a kinase, phosphatase, protease, protease inhibitor,  
 CC transporter, cytoskeletal protein, receptor or transcription factor. The  
 CC polypeptide is useful as an immunogenic composition for eliciting in a  
 CC mammal an immunogenic response directed against any of the purified  
 CC polypeptide. The purified polypeptide, or the antibody that binds to this  
 CC polypeptide, is useful for treating cancer. The polypeptide is also  
 CC useful for identifying compounds that binds to a naturally processed  
 CC class I or class II MHC-binding polypeptide. The polypeptides and  
 CC polynucleotides are particularly useful for treating or preventing  
 CC myeloma, colon cancer, gastric cancer, adenocarcinoma, sarcoma, melanoma,  
 CC lymphoma or leukemia. These are also useful for screening agents for  
 CC treating the above mentioned diseases. This sequence represents an  
 CC expressed protein tag (EPT) isolated from human tissue for translational  
 CC profiling. Note: This sequence does not appear in the printed  
 CC specification but was obtained in electronic format directly from WIPO at  
 CC ftp.wipo.int/pub/published\_pct\_sequences  
 XX Sequence 524 AA;  
 SO  
 Query Match 100.0%; Score 2789; DB 6; Length 524;  
 Best Local Similarity 100.0%; Pred. No. 2.3e-237;  
 Matches 524; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 MYALFLASLLGALAGPVLGLKECTGSAWCONVTASDCGAKHCLQTVNNKPTVKS 60  
 DB 1 MYALFLASLLGALAGPVLGLKECTGSAWCONVTASDCGAKHCLQTVNNKPTVKS 60  
 QY 61 LPDCICKDVYTAAGDMKDNATEEEILVLEKTCDMLPKPNMSASCKEIVDSYLPVLDI 120  
 DB 61 LPDCICKDVYTAAGDMKDNATEEEILVLEKTCDMLPKPNMSASCKEIVDSYLPVLDI 120  
 QY 121 IKGEMSRGECVCSALNCEISLOKHLAELNOKOLESNKIPELDMTEVAPFMANIPLLLY 180  
 DB 121 IKGEMSRGECVCSALNCEISLOKHLAELNOKOLESNKIPELDMTEVAPFMANIPLLLY 180  
 QY 181 PDGPRSKPDKNGDVCDQCIQWVTDIQTAVRTNSTFVQALVEHYKECDRLGPGMADI 240  
 DB 181 PDGPRSKPDKNGDVCDQCIQWVTDIQTAVRTNSTFVQALVEHYKECDRLGPGMADI 240

DB 181 PDGPRSKPDKNGDVCDQCIQWVTDIQTAVRTNSTFVQALVEHYKECDRLGPGMADI 240  
 QY 241 CKNTISOYSEIAIOMMMHMPKEICALVGFCDVKEKMPQTLVPAKASKNVLPALVELVE 300  
 DB 241 CKNTISOYSEIAIOMMMHMPKEICALVGFCDVKEKMPQTLVPAKASKNVLPALVELVE 300  
 QY 301 PIKHEVPASDVYCEVCEFLVKEVTKLIDNNKTEKEILDAFDMCSKLPKSLSEECQEV 360  
 DB 301 PIKHEVPASDVYCEVCEFLVKEVTKLIDNNKTEKEILDAFDMCSKLPKSLSEECQEV 360  
 QY 361 VDTGSSILSLLEVSPELVCSMLHCSTRLPALVHTQXDGDFCEVCKKLVGYLD 420  
 DB 361 VDTGSSILSLLEVSPELVCSMLHCSTRLPALVHTQXDGDFCEVCKKLVGYLD 420  
 QY 421 RNLEKSTKQELIALAEGKCSFLPDYQKQDQFVAEYEPVLJEILVEWMDPSFVCLKIG 480  
 DB 421 RNLEKSTKQELIALAEGKCSFLPDYQKQDQFVAEYEPVLJEILVEWMDPSFVCLKIG 480  
 QY 481 ACPSAHKPLLGTEKCIWGPSTWCONTEYTAACNAVEHCKRHVNN 524  
 DB 481 ACPSAHKPLLGTEKCIWGPSTWCONTEYTAACNAVEHCKRHVNN 524

Search completed: May 5, 2004, 13:31:03  
 Job time : 75.151 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 5, 2004, 13:41:29 ; Search time 38.2188 Seconds  
(without alignments)  
4325.921 Million cell updates/sec

Title: US-09-743-684A-1

Perfect score: 524

Sequence: 1 MYALFLSLGALAGPVL.....NTETAACNAVEHCKRHVN 524

Scoring table: OLIGO  
Gapop 60.0 , Gapext 60.0

Searched: 1017041 seqs, 315518202 residues

Word size : 0

Total number of hits satisfying chosen parameters: 1017041

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : SPTREMBL 25.\*

- 1: sp\_archaea.\*
- 2: sp\_bacteria.\*
- 3: sp\_fungi.\*
- 4: sp\_human.\*
- 5: sp\_invertebrate.\*
- 6: sp\_mammal.\*
- 7: sp\_mhc.\*
- 8: sp\_organelle.\*
- 9: sp\_phase.\*
- 10: sp\_plant.\*
- 11: sp\_rodent.\*
- 12: sp\_virus.\*
- 13: sp\_vertebrate.\*
- 14: sp\_unclassified.\*
- 15: sp\_rvirus.\*
- 16: sp\_bacteriap.\*
- 17: sp\_archaeap.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	22	4.2	554	11 Q8BFQ1	Q8bfq1 mus musculus
2	20	3.8	121	6 P79254	P79254 ovis aries
3	13	2.5	13	4 O75905	O75905 homo sapien
4	13	2.5	512	13 Q7S70	Q7sy70 xenopus lae
5	12	2.3	449	11 Q8BJV5	Q8bjv5 mus musculus
6	12	2.3	525	11 Q8C1C1	Q8clc1 mus musculus
7	11	2.1	402	11 Q8C1N0	Q8cln0 mus musculus
8	9	1.7	520	13 Q8UWZ4	Q8uwz4 brachydanio
9	9	1.7	522	13 Q9DG82	Q9dgs2 brachydanio
10	8	1.5	50	9 Q854G1	Q854g1 mycobacteri
11	8	1.5	90	9 Q9ZXL7	Q9zxl7 bacterioph
12	8	1.5	154	9 O21999	O21999 bacterioph
13	8	1.5	162	16 Q928C7	Q928c7 listeria in
14	8	1.5	165	16 Q986B3	Q986b3 rhizobium 1
15	8	1.5	177	4 Q9NSM6	Q9nsm6 homo sapien
16	8	1.5	179	2 Q47224	Q47224 escherichia

17	8	1.5	179	16 Q8CVG2	Q8cvg2 escherichia
18	8	1.5	187	16 Q82W66	Q82w66 nitrosomona
19	8	1.5	188	2 Q84GH7	Q84gh7 gamma-prote
20	8	1.5	201	2 Q44889	Q44889 bordetella
21	8	1.5	201	16 Q7WGY5	Q7wgy5 bordetella
22	8	1.5	212	16 Q7UAQ7	Q7uaq7 shigella fl
23	8	1.5	215	16 Q8XC62	Q8xc62 escherichia
24	8	1.5	215	16 Q83P69	Q83p69 shigella fl
25	8	1.5	216	10 Q9LYQ4	Q9lyq4 arabidopsis
26	8	1.5	225	10 Q48998	Q48998 oryza sativ
27	8	1.5	294	13 Q98S10	Q98s10 brachydanio
28	8	1.5	297	16 Q7VFN8	Q7vpfn8 haemophilus
29	8	1.5	300	2 Q93UN6	Q93un6 vibrio vuln
30	8	1.5	300	16 Q8D492	Q8d492 vibrio vuln
31	8	1.5	326	16 Q8X584	Q8x584 escherichia
32	8	1.5	326	16 Q83RP7	Q83rp7 shigella fl
33	8	1.5	330	2 Q9X7G2	Q9x7g2 methylobact
34	8	1.5	337	12 Q8JTY7	Q8jty7 lumpy skin
35	8	1.5	337	12 Q8JTN0	Q8jtn0 lumpy skin
36	8	1.5	337	12 Q91ML5	Q91ml5 lumpy skin
37	8	1.5	354	16 Q8YBQ5	Q8ybg5 brucella me
38	8	1.5	373	2 Q937R1	Q937r1 brucella me
39	8	1.5	375	16 Q8FWM3	Q8fwm3 brucella su
40	8	1.5	377	2 Q8L243	Q8l243 salmonella
41	8	1.5	377	2 Q84922	Q84922 escherichia
42	8	1.5	377	2 Q7WTI0	Q7wti0 escherichia
43	8	1.5	379	2 Q9EXN9	Q9exn9 escherichia
44	8	1.5	390	5 Q962K0	Q962k0 plasmodium
45	8	1.5	391	2 Q9F170	Q9f170 salmonella

#### ALIGNMENTS

#### RESULT 1

ID	Q8BFQ1	PRELIMINARY;	PRT;	554 AA.
AC	Q8BFQ1			
DT	01-MAR-2003 (TrEMBLrel. 23, Created)			
DT	01-MAR-2003 (TrEMBLrel. 23, Last sequence update)			
DT	01-OCT-2003 (TrEMBLrel. 25, Last annotation update)			
DE	Prosaposin.			
GN	PSAP.			
OS	Mus musculus (Mouse).			
OC	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.			
OX	NCBI_TaxID=10090;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN=NOD; TISSUE=Kidney, and Thymus;			
RX	MEDLINE=22354683; PubMed=12466851;			
RA	The FANTOM Consortium,			
RA	the RIKEN Genome Exploration Research Group Phase I & II Team;			
RT	"Analysis of the mouse transcriptome based on functional annotation of			
RL	Nature 420:563-573(2002)."			
DR	EMBL; AK088369; BAC40308.1; -			
DR	MGD; MGI:97783; Psap.			
DR	GO; GO:0005764; C:lysosome; IEA.			
DR	GO; GO:0006665; P:sphingolipid metabolism; IEA.			
DR	InterPro; IPR003119; SapA.			
DR	InterPro; IPR007856; SapB_1.			
DR	InterPro; IPR008138; SapB_2.			
DR	InterPro; IPR008140; SapB_sub.			
DR	InterPro; IPR008373; Saposin.			
DR	Pfam; PF02199; SAPA; 2.			
DR	Pfam; PF05184; SapB_1; 4.			
DR	Pfam; PF03489; SapB_2; 4.			
DR	PRINTS; PR01797; SAPOSIN.			
DR	ProDom; PD001732; SapB_sub; 3.			
DR	SMART; SM00162; SAPA; 2.			

```
DR SMART; SM00118; SAPB; 4.
SQ SEQUENCE 554 AA; 61050 MW; FF58DB79C7CC0C18 CRC64;

Query Match
Best Local Similarity 4.2%; Score 22; DB 11; Length 554;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 431 ETLAALKGCSELPDPYKQCD 452
Db 461 ETLAALKGCSELPDPYKQCD 482

RESULT 2
ID P79254 PRELIMINARY; PRT; 121 AA.
AC P79254;
DT 01-MAY-1997 (TREMBLrel. 03, Created)
DT 01-MAY-1997 (TREMBLrel. 03, Last sequence update)
DE Sulfated glycoprotein-1/SGP-1 (Fragment).
OS Ovis aries (Sheep).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovidae; Caprinae; Ovis.
OX NCBI_TaxID=9940;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=96249303; PubMed=8848570;
RA Spencer T.E., Graf G.H., Bazer F.W.;
RT "Sulfated glycoprotein-1 (SGP-1) expression in ovine endometrium
during the oestrous cycle and early pregnancy.";
RL Reprod. Fert. Dev. 7:1053-1060(1995).
DR EMBL; S82555; AAD14405.1; -.
DR GO; GO:0005764; C:lysosome; IEA.
DR GO; GO:0006665; P:sphingolipid metabolism; IEA.
DR InterPro; IPR008138; SapB_2.
DR InterPro; IPR008373; Saposin.
DR Pfam; PF03489; SapB_2; 1.
DR PRINTS; PR01797; SAPOSIN.
DR SMART; SM00118; SAPB; 1.
FT NON TER 1
SQ SEQUENCE 121 AA; 13604 MW; 4F0F5A6EB83D0C9A CRC64;

Query Match
Best Local Similarity 3.8%; Score 20; DB 6; Length 121;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 225 HVKECDRLGPGMADICKNY 244
Db 21 HVKECDRLGPGMADICKNY 40

RESULT 3
ID P79254 PRELIMINARY; PRT; 13 AA.
AC P79254;
DT 01-MAR-2002 (TREMBLrel. 08, Created)
DT 01-MAR-2002 (TREMBLrel. 20, Last sequence update)
DE Prosaposin (Fragment).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=98426155; PubMed=9751800;
RA Sun Y., Jin P., Witte D.P., Grabowski G.A.;
RT "Isolation and characterization of the human prosaposin promoter.";
RL Gene 218:37-47(1998).
FT NON TER 1
SQ SEQUENCE FROM N.A.
```

```
RA Sun Y., Jin P., Grabowski G.A.;
RT "Isolation and characterization of the human prosaposin promoter.";
RL Gene 218:37-47(1998).
RN [3]
RP SEQUENCE FROM N.A.
RX MEDLINE=21205763; PubMed=11309366;
RA Hukova H., Cervenkova M., Ledvinov J., Tochackova M., Hrebicek M.,
RA Poupetova H., Befekadu A., Berna L., Paton B.C., Harzer K., Boeozar A.,
RA Smid F., Ellender M.;
RT "A novel mutation in the coding region of the prosaposin gene leads to
a complete deficiency of prosaposin and saposins, and is associated
with a complex sphingolipidosis dominated by lactosylceramide
accumulation.";
RL Hum. Mol. Genet. 10:927-940(2001).
DR EMBL; AF057307; AAC64921.2; -.
DR NON TER 13
FT NON TER 13
SQ SEQUENCE 13 AA; 1383 MW; 57585B6C1784D727 CRC64;

Query Match
Best Local Similarity 2.5%; Score 13; DB 4; Length 13;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MYALFLLASLLGA 13
Db 1 MYALFLLASLLGA 13

RESULT 4
ID Q7SV70 PRELIMINARY; PRT; 512 AA.
AC Q7SV70;
DT 01-OCT-2003 (TREMBLrel. 25, Created)
DT 01-OCT-2003 (TREMBLrel. 25, Last sequence update)
DE Hypothetical protein.
OS Xenopus laevis (African clawed frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae;
OC Xenopodinae; Xenopus.
OX NCBI_TaxID=8355;
RN [1]
RP SEQUENCE FROM N.A.
RX TISSUE=Whole;
RT MEDLINE=223441132; PubMed=12454917;
RA Klein S.L., Strausberg R.L., Wagner L., Pontius J., Clifton S.W.,
RA Richardson P.;
RT "Genetic and genomic tools for Xenopus research: The NIH Xenopus
RT initiative.";
RL Dev. Dyn. 225:384-391(2002).
RN [2]
RP SEQUENCE FROM N.A.
RX TISSUE=Whole;
RT MEDLINE=22388257; PubMed=12477932;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Heide F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Ustin T.B., Toshiyuki S., Carninci P., Prange C.J.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Pahey J., Helton E., Kettman M., Madan A., Rodrigues S., Sanchez A.,
RA Whiting R., Touchman J.W., Green E.D., Dickson M.C.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA Krzywinski M.I., Skalska U., Smallos D.E., Schnerch A., Schein J.E.,
RA Jones S.J., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences.";
```

RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).

RN [3]

RP SEQUENCE FROM N.A.

RC TISSUE=Whole;

RA Klein S., Strauberg R., to the EMBL/GenBank/DBJ databases.

RL Submitted (JUL-2003)

DR EMBL; BC054988; AAH54988.1; -

KW Hypothetical protein.

SQ SEQUENCE 512 AA; 57555 MW; 57CFASE2093F6FB1 CRC64;

Query Match

Best Local Similarity 2.5%; Score 13; DB 13; Length 512;

Matches 13; Conservative. 0; Mismatches 0; Indels 0; Gaps 0;

Qy 52 VWNKPTVKSLPCD 64

Db 52 VWNKPTVKSLPCD 64

RESULT 5

Q8BJV5

ID Q8BJV5

AC Q8BJV5

DT 01-MAR-2003

DT 01-MAR-2003

DT 01-OCT-2003

DE Hypothetical saposin A-type domain/saposin type B containing

DE protein.

GN 2310020A21RIK.

OS Mus musculus (Mouse).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

OX NCBI\_TaxID=10090;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=C57BL/6J; TISSUE=Eye;

RX MEDLINE=22354683; PubMed=12466851;

RA The FANTOM Consortium,

RA the RIKEN Genome Exploration Research Group Phase I & II Team;

RT "Analysis of the mouse transcriptome based on functional annotation of

RT 60,770 full-length cDNAs."

RL Nature 420:563-573 (2002).

DR EMBL; AK078699; BAC37363.1; -

DR MGD; MGI:1924193; 2310020A21RIK.

DR GO; GO:0005764; C:lysosome; IEA.

DR GO; GO:0006665; P:sphingolipid metabolism; IEA.

DR InterPro; IPR007856; SapB\_1.

DR InterPro; IPR007856; SapB\_1.

DR InterPro; IPR008138; SapB\_sub.

DR InterPro; IPR008138; SapB\_sub.

DR InterPro; IPR008373; Saposin.

DR Pfam; PF02199; SAPA; 1.

DR Pfam; PF05184; SapB\_1; 3.

DR PRINTS; PR01797; SAPOSIN.

DR PRODOM; PD001732; SapB\_sub; 2.

DR SMART; SM00162; SAPA; 1.

DR SMART; SM00118; SAPP; 3.

KW Hypothetical protein.

SQ SEQUENCE 449 AA; 48705 MW; 377CC42A475B292F CRC64;

Query Match

Best Local Similarity 2.3%; Score 12; DB 11; Length 449;

Matches 12; Conservative. 0; Mismatches 0; Indels 0; Gaps 0;

Qy 52 VWNKPTVKSLPC 63

Db 56 VWNKPTVKSLPC 67

RESULT 6

Q8C1C1

ID Q8C1C1

AC Q8C1C1

DT 01-MAR-2003

DT 01-MAR-2003

DT 01-OCT-2003

DE Hypothetical saposin A-type domain/saposin type B containing

DE protein.

GN 2310020A21RIK.

OS Mus musculus (Mouse).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

OX NCBI\_TaxID=10090;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=C57BL/6J; TISSUE=Tongue;

RX MEDLINE=22354683; PubMed=12466851;

RA The FANTOM Consortium,

RA the RIKEN Genome Exploration Research Group Phase I & II Team;

RT "Analysis of the mouse transcriptome based on functional annotation of

RT 60,770 full-length cDNAs."

RL Nature 420:563-573 (2002).

DR EMBL; AK009408; BAC25258.1; --  
 DR MGD; MGI:1924193; 2310020A21Rik.  
 DR GO; GO:0005764; C.lysosome; IEA.  
 DR GO; GO:0006665; P.sphingolipid metabolism; IEA.  
 DR InterPro; IPR003119; Sapa.  
 DR InterPro; IPR007856; SapaB.1.  
 DR InterPro; IPR008138; SapaB.2.  
 DR InterPro; IPR008140; SapaB.sub.  
 DR InterPro; IPR008373; Saposin.  
 DR InterPro; IPR008139; SaposinB.  
 DR Pfam; PF02199; SAPA; 1.  
 DR Pfam; PF05184; SapaB\_1; 3.  
 DR Pfam; PF03489; SapaB\_2; 3.  
 DR PRINTS; PR01797; SAPOSIN.  
 DR ProDom; PD001732; SapaB.sub; 2.  
 DR SMART; SM00162; SAPA; 1.  
 DR SMART; SM00118; SapaB; 3.  
 KW Hypothetical protein.  
 SQ SEQUENCE 402 AA; 44420 MW; E90017CBF4017ED6 CRC64;

Query Match 2.1%; Score 11; DB 11; Length 402;  
 Best Local Similarity 100.0%; Pred. No. 0.065;  
 Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 454 FVAEYEPVLIE 464  
 DB 322 FVAEYEPVLIE 332

RESULT 8  
 Q8UVZ4 PRELIMINARY; PRT; 520 AA.  
 ID Q8UVZ4  
 AC Q8UVZ4  
 DT 01-MAR-2002 (TrEMBLrel. 20, Created)  
 DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)  
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
 DE Lysosomal cofactor/neurotrophic factor prosaposin.  
 OS Brachydanio rerio (Zebrafish) (Danio rerio).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;  
 OC Cyprinidae; Danio.  
 ON NCBI\_TaxID=7955;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Seo H.-C., Lie O., Fjose A., O'Brien J.S., Kishimoto Y.,  
 RA "Cloning, expression and promoter analysis of zebrafish prosaposin.";  
 RT Submitted (JUN-2000) to the EMBL/GenBank/DBJ databases.  
 RL EMBL; AF276996; AAL54381.1; --  
 DR GO; GO:0005764; C.lysosome; IEA.  
 DR GO; GO:0006665; P.sphingolipid metabolism; IEA.  
 DR InterPro; IPR003119; Sapa.  
 DR InterPro; IPR007856; SapaB.1.  
 DR InterPro; IPR008138; SapaB.2.  
 DR InterPro; IPR008140; SapaB.sub.  
 DR InterPro; IPR008373; Saposin.  
 DR InterPro; IPR008139; SaposinB.  
 DR Pfam; PF02199; SAPA; 2.  
 DR Pfam; PF05184; SapaB\_1; 3.  
 DR Pfam; PF03489; SapaB\_2; 4.  
 DR PRINTS; PR01797; SAPOSIN.  
 DR ProDom; PD001732; SapaB.sub; 3.  
 DR SMART; SM00162; SAPA; 2.  
 DR SMART; SM00118; SapaB; 4.  
 SQ SEQUENCE 520 AA; 57431 MW; F9E20F84BA41CB5 CRC64;

Query Match 1.7%; Score 9; DB 13; Length 520;  
 Best Local Similarity 100.0%; Pred. No. 7.8;  
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 32 WCONVKTAS 40  
 DB 30 WCONVKTAS 38

RESULT 9  
 Q9DG82 PRELIMINARY; PRT; 522 AA.  
 ID Q9DG82  
 AC Q9DG82  
 DT 01-MAR-2001 (TrEMBLrel. 16, Created)  
 DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)  
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
 DE Prosaposin.  
 GN PSAP.  
 OS Brachydanio rerio (Zebrafish) (Danio rerio).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;  
 OC Cyprinidae; Danio.  
 ON NCBI\_TaxID=7955;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Altman N., Horowitz M.;  
 RT "The zebrafish prosaposin cDNA.";  
 RL Submitted (NOV-1998) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AF108655; AAG32919.1; --  
 DR ZFIN; ZDB-GENE-020108-1; psap.  
 DR GO; GO:0005764; C.lysosome; IEA.  
 DR GO; GO:0006665; P.sphingolipid metabolism; IEA.  
 DR InterPro; IPR003119; Sapa.  
 DR InterPro; IPR007856; SapaB.1.  
 DR InterPro; IPR008138; SapaB.2.  
 DR InterPro; IPR008140; SapaB.sub.  
 DR InterPro; IPR008373; Saposin.  
 DR InterPro; IPR008139; SaposinB.  
 DR Pfam; PF02199; SAPA; 2.  
 DR Pfam; PF05184; SapaB\_1; 3.  
 DR Pfam; PF03489; SapaB\_2; 4.  
 DR PRINTS; PR01797; SAPOSIN.  
 DR ProDom; PD001732; SapaB.sub; 3.  
 DR SMART; SM00162; SAPA; 2.  
 DR SMART; SM00118; SapaB; 4.  
 SQ SEQUENCE 522 AA; 57671 MW; D3C15A305725C1CD CRC64;

Query Match 1.7%; Score 9; DB 13; Length 522;  
 Best Local Similarity 100.0%; Pred. No. 7.9;  
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 32 WCONVKTAS 40  
 DB 30 WCONVKTAS 38

RESULT 10  
 Q854G1 PRELIMINARY; PRT; 50 AA.  
 ID Q854G1  
 AC Q854G1  
 DT 01-JUN-2003 (TrEMBLrel. 24, Created)  
 DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)  
 DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)  
 DE Gp106.  
 GN Gp106.  
 OS Mycobacteriophage Omega.  
 OC Viruses; dsDNA viruses, no RNA stage; Caudovirales; Siphoviridae.  
 ON NCBI\_TaxID=205879;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=22592660; PubMed=12705866;  
 RA Pedulla M.L., Ford M.E., Houtz J.M., Karthikeyan T., Wadsworth C.,  
 RA Lewis J.A., Jacobs-Sera D., Falbo J., Gross J., Pannunzio N.R.,  
 RA Brucker W., Kumar V., Kandasamy J., Keenan L., Bardarov S.,  
 RA Kriakov J., Lawrence J.G., Jacobs W.R. Jr., Hendrix R.W.,  
 RA Hatfull G.F.;  
 RT "Origins of highly mosaic mycobacteriophage genomes.";  
 RL Cell 113:171-182(2003).  
 DR EMBL; AY129338; AAN12747.1; --  
 SQ SEQUENCE 50 AA; 5689 MW; 16993D01C4E247F4 CRC64;

```

Query Match          1.5%; Score 8; DB 9; Length 50;
Best Local Similarity 100.0%; Pred. No. 11;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 117 ILDIKGE 124
Db 43 ILDIKGE 50

RESULT 11
Q9ZXL7 PRELIMINARY; PRT; 90 AA.
AC Q9ZXL7;
DT 01-MAY-1999 (TrEMBLrel. 10, Created)
DT 01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Orl0.
OS Bacteriophage phi CTX.
OC Viruses; dsDNA viruses, no RNA stage; Caudovirales; Myoviridae.
OC NCBI_TaxID=35343;
RN [1]
RP STRAIN=phiCTX-c;
RA Hayashi T.;
RL Submitted (OCT-1997) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RP STRAIN=phiCTX-c;
RX MEDLINE=99157549; PubMed=10027959;
RA Nakayama K., Kanaya S., Ohnishi M., Terawaki Y., Hayashi T.;
RT "The complete nucleotide sequence of phiCTX, a cytotoxin-converting
RT phage of Pseudomonas aeruginosa: implications for phage evolution and
RT horizontal gene transfer via bacteriophage.";
RL Mol. Microbiol. 31:399-419(1999).
RN [3]
RP SEQUENCE FROM N.A.
RP STRAIN=phiCTX-c;
RX MEDLINE=90014160; PubMed=2507866;
RA Hayashi T., Kamio Y., Hishinuma F., Usami Y., Titani K., Terawaki Y.;
RT "Pseudomonas aeruginosa cytotoxin: the nucleotide sequence of the gene
RT and the mechanism of activation of the protoxin.";
RL Mol. Microbiol. 3:861-868(1989).
DR EMBL; AB008550; BAA36236.1; -.
DR InterPro; IPR008473; DUF754.
DR Pfam; PF05449; DUF754; 1.
SQ SEQUENCE 90 AA; 9698 MW; 7E8C73C84E24F4D4 CRC64;

Query Match          1.5%; Score 8; DB 9; Length 90;
Best Local Similarity 100.0%; Pred. No. 17;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 6 LLASLLGA 13
Db 37 LLASLLGA 44

RESULT 12
O21999 PRELIMINARY; PRT; 154 AA.
AC O21999;
DT 01-JAN-1998 (TrEMBLrel. 05, Created)
DT 01-JAN-1998 (TrEMBLrel. 05, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE LacZ-alpha.
OS Bacteriophage lambda.
OG Plasmid pAL-2.
OC Viruses; dsDNA viruses, no RNA stage; Caudovirales; Siphoviridae;
OC Lambda-like viruses.
OC NCBI_TaxID=10710;
RN [1]
RP SEQUENCE FROM N.A.
RP TRANSPOSON=IS1;
RX MEDLINE=97473533; PubMed=9332387;

```

```

RA Ahmed A., Podemski L.;
RT "Use of ordered deletions in genome sequencing.";
RL Gene 197:367-373(1997).
DR EMBL; U66885; AAC48874.1; -.
DR HSP; P00722; IBGL.
DR GO; GO:0046821; C:extrachromosomal DNA; IEA.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0004553; F:hydrolase activity, hydrolyzing O-glycosyl . . .; IEA.
DR GO; GO:0015520; F:tetracycline:hydrogen antiporter activity; IEA.
DR GO; GO:0005215; F:transporter activity; IEA.
DR GO; GO:0005975; P:carbohydrate metabolism; IEA.
DR GO; GO:0015904; P:tetracycline transport; IEA.
DR GO; GO:0006810; P:transport; IEA.
DR InterPro; IPR008979; Gal bind like.
DR InterPro; IPR006104; Glyco hydro 2SB.
DR InterPro; IPR005829; Sug transporter.
DR InterPro; IPR001958; TCR_TetA.
DR Pfam; PF02837; Glyco hydro 2_N; 1.
DR PRINTS; PR01035; TCRTETA.
DR PROSITE; PS00216; SUGAR_TRANSPORT_1; 1.
KW Plasmid.
SQ SEQUENCE 154 AA; 17242 MW; 2CEE57E8FA128370 CRC64;

Query Match          1.5%; Score 8; DB 9; Length 154;
Best Local Similarity 100.0%; Pred. No. 27;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 6 LLASLLGA 13
Db 123 LLASLLGA 130

RESULT 13
Q928C7 PRELIMINARY; PRT; 162 AA.
AC Q928C7;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Hypothetical protein lin2609.
GN Lin2609.
OS Listeria innocua.
OC Bacteria; Firmicutes; Bacillales; Listeriaceae; Listeria.
OC NCBI_TaxID=1642;
RN [1]
RP SEQUENCE FROM N.A.
RP STRAIN=CLIP 11262 / Serovar 6a;
RX MEDLINE=21537279; PubMed=11679669;
RA Glaser P., Frangeul L., Buchrieser C., Rusniok C., Amend A.,
RA Baquero F., Berche P., Bloeker H., Brandt P., Chakraborty T.,
RA Charbit A., Chetouani F., Couve E., de Daruvar A., Dehoux P.,
RA Domann E., Dominguez-Bernal G., Duchaud E., Durant L., Dussurget O.,
RA Etian K.-D., Fsihi H., Garcia-del Portillo F., Garrido P.,
RA Gautier L., Goebel W., Gomez-Lopez N., Hain T., Hauf J., Jackson D.,
RA Jones L.-M., Kaerst U., Krefit J., Kuhn M., Kunst F., Kurapat G.,
RA Madueno E., Maitournam A., Mata Vicente J., Ng E., Nedjari H.,
RA Nordsiek G., Novella S., de Pablo B., Perez-Diaz J.-C., Purcell R.,
RA Rammel B., Rose M., Schlueter T., Simoes N., Tierrez A.,
RA Vazquez-Boland J.-A., Vose H., Wehland J., Cossart P.;
RT "Comparative genomics of Listeria species.";
RL Science 294:849-852(2001).
DR EMBL; AL596173; CAC97836.1; -.
DR PIR; AD1758; AD1758.
DR Listlist; LIN02609; -.
DR InterPro; IPR000437; Prok lipoprot S.
DR PROSITE; PS00013; PROKAR_LIPOPROTEIN; 1.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 162 AA; 18193 MW; 181DA681AE002BC7 CRC64;

Query Match          1.5%; Score 8; DB 16; Length 162;
Best Local Similarity 100.0%; Pred. No. 29;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY 4 LFLASLL 11  
| | | | |  
Db 4 LFLASLL 11

Db 146 EILAALEK 153

Search completed: May 5, 2004, 13:47:43  
Job time : 39.2188 secs

## RESULT 14

Q986B3 ID Q986B3 PRELIMINARY; PRT; 165 AA.  
AC Q986B3;  
DT 01-OCT-2001 (TrEMBLrel. 18, Created)  
DT 01-OCT-2001 (TrEMBLrel. 18, Last sequence update)  
DE 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)  
DE Hypothetical protein mll17437.  
GN MLL7437.  
OS Rhizobium loti (Mesorhizobium loti).  
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;  
OC Phyllobacteriaceae; Mesorhizobium.  
OX NCBI\_TaxID=381;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=MAFF303099;  
RX MEDLINE=21082930; PubMed=11214968;  
RA Kaneko T., Nakamura Y., Sato S., Asamizu E., Kato T., Sasamoto S.,  
RA Watanabe A., Ideawara K., Ishikawa A., Kawashima K., Kimura T.,  
RA Kishida Y., Kiyokawa C., Kohara M., Matsumoto M., Matsuno A.,  
RA Mochizuki Y., Nakayama S., Nakazaki N., Shimo S., Sugimoto M.,  
RA Takeuchi C., Yamada M., Tabata S.;  
RT "Complete genome structure of the nitrogen-fixing symbiotic bacterium  
RT Mesorhizobium loti.";  
RL DNA Res. 7:331-338(2000).  
DR EMBL; AP003011; BAB53540.1; -.  
KW Hypothetical protein; Complete proteome.  
SQ SEQUENCE 165 AA; 18468 MW; B8FA62EEAA271361 CRC64;

Query Match 1.5%; Score 8; DB 16; Length 165;  
Best Local Similarity 100.0%; Pred. No. 29;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 6 LLASLLGA 13  
| | | | |  
Db 49 LLASLLGA 56

## RESULT 15

Q9NSM6 ID Q9NSM6 PRELIMINARY; PRT; 177 AA.  
AC Q9NSM6;  
DT 01-OCT-2000 (TrEMBLrel. 15, Created)  
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)  
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)  
DE Hypothetical protein (Fragment).  
GN DKFZP761F1921.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Amalgam;  
RA Poustka A., Wellenreuther R., Mewes H.W., Weil B., Wiemann S.;  
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AL162013; CAB82365.1; -.  
DR PIR; T47133; T47133.  
KW Hypothetical protein.  
FT NON TER 1  
SQ SEQUENCE 177 AA; 20522 MW; EB717DE9993AF417 CRC64;

Query Match 1.5%; Score 8; DB 4; Length 177;  
Best Local Similarity 100.0%; Pred. No. 31;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 431 EILAALEK 438  
| | | | |



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 5, 2004, 13:35:24 ; Search time 10.5605 Seconds  
(without alignments)  
2583.672 Million cell updates/sec

Title: US-09-743-684A-1

Perfect score: 524

Sequence: 1 MYALFLIALSLGALAGFLV.....NTETAACQNAVEHCKRHVMN 524

Scoring table: OLIGO

Gapop 60.0 , Gapext 60.0

Searched: 141681 seqs, 52070155 residues

Word size : 0

Total number of hits satisfying chosen parameters: 141681

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : SwissProt\_42.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Query Match	Score	Length	ID	Description
1	524	100.0	524	1	SAP HUMAN
2	30	5.7	525	1	SAP_BOVIN
3	22	4.2	80	1	SAP_PIG
4	22	4.2	554	1	SAP_RAT
5	22	4.2	557	1	SAP_MOUSE
6	12	2.3	518	1	SAP_CHICK
7	9	1.7	81	1	SAP_CAVPO
8	8	1.5	179	1	FIMI_ECOLI
9	8	1.5	186	1	HS23_DROME
10	8	1.5	201	1	FMFX_BORPE
11	8	1.5	216	1	Y726_METJA
12	8	1.5	349	1	BOR4_MOUSE
13	8	1.5	356	1	BOR4_HUMAN
14	8	1.5	396	1	TCR3_ECOLI
15	8	1.5	435	1	MNTC_BACSU
16	8	1.5	490	1	FA10_RABIT
17	8	1.5	655	1	YHCP_ECOLI
18	8	1.5	813	1	FCR2_XENLA
19	7	1.3	84	1	ACPX_STRAW
20	7	1.3	105	1	RN37_YEAST
21	7	1.3	112	1	MTP_CHICK
22	7	1.3	137	1	CC23_HOMAM
23	7	1.3	144	1	Y638_TREPA
24	7	1.3	145	1	PTMA_BACST
25	7	1.3	158	1	ILVH_BUCAI
26	7	1.3	159	1	ILVH_BUCBP
27	7	1.3	164	1	ID1_RAT
28	7	1.3	168	1	ID1_MOUSE
29	7	1.3	174	1	NU6M_HUMAN
30	7	1.3	203	1	SODM_THETH
31	7	1.3	208	1	MOBA_RHILIO
32	7	1.3	210	1	AFS1_SCHPO
33	7	1.3	210	1	Y598_SCHPO

#### RESULT 1

ID	SAP_HUMAN	STANDARD;	PRT;	524 AA.
AC	P07602; P07292; P15793; P78538; P78541; P78546; P78547; P78558;			
AC	Q92739; Q92740; Q92741; Q92742;			
DT	01-APR-1988 (Rel. 07, Created)			
DT	01-APR-1990 (Rel. 14, Last sequence update)			
DT	10-OCT-2003 (Rel. 42, Last annotation update)			
DE	Proactivator polypeptide precursor [Contains: Saposin A (Protein A); Saposin B (Sphingolipid activator protein 1) (SAP-1) (Cerebroside sulfate activator) (CSAct) (Dispersin) (Sulfatide/GMI activator); Saposin C (Co-beta-glucosidase) (Al activator) (Glucosylceramidase activator) (Sphingolipid activator protein 2) (SAP-2); Saposin D (Protein C) (Component C)].			
GN	PSAP.			
OS	Homo sapiens (Human).			
OC	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.			
OX	NCBI_TaxID=9606;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	TISSUE=Liver;			
RX	MEDLINE=90129043; PubMed=2515150;			
RA	Rorman E.G., Grabowski G.A.;			
RT	"Molecular cloning of a human co-beta-glucosidase cDNA: evidence that four sphingolipid hydrolase activator proteins are encoded by single genes in humans and rats.";			
RL	Genomics 5:486-492(1989).			
RN	[2]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=89255151; PubMed=2498298;			
RA	Nakano T., Sandhoff K., Stuenkel J., Christomanou H., Suzuki K.;			
RT	"Structure of full-length cDNA coding for sulfatide activator, a Co-beta-glucosidase and two other homologous proteins: two alternate forms of the sulfatide activator.";			
RL	J. Biochem. 105:152-154(1989).			
RN	[3]			
RP	SEQUENCE FROM N.A. (ISOFORM SAP-MU-0).			
RC	TISSUE=Brain, Eye, and Skin;			
RX	MEDLINE=22388257; PubMed=12477932;			
RA	Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G., Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D., Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K., Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F., Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L., Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E., Brownstein M.J., Udén T.B., Toshiyuki S., Carninci P., Prange C., Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullany S.J., Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H., Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W., Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A., Fahey J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A., Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G., Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C., Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,			

P74368 synechocyst  
P23958 phaseolus v  
P42270 escherichia  
P15764 cyanophora  
P77395 escherichia  
Q9hxq6 pseudomonas  
Q88mvo pseudomonas  
Q9hvz5 pseudomonas  
P31118 paramecium  
Q9juu2 neisseria m  
Q9jz80 neisseria m

#### ALIGNMENTS

RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;  
RT "Generation and initial analysis of more than 15,000 full-length  
RT human and mouse cDNA sequences.";  
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).  
RN [4]  
RP SEQUENCE OF 59-125 AND 304-513 FROM N.A.  
RC TISSUE=Brain;  
RX MEDLINE=91192146; PubMed=2013321;  
RA Holtschmidt H., Sandhoff K., Fuerst W., Kwon H.Y., Schnabel D.,  
RA Suzuki K.;  
RT "The organization of the gene for the human cerebroside sulfate  
RT activator protein.";  
RL FEBS Lett. 280:267-270 (1991).  
RN [5]  
RP SEQUENCE OF 164-524 FROM N.A.  
RX MEDLINE=88068647; PubMed=2825202;  
RA Dewji N.N., Wenger D.A., O'Brien J.S.;  
RT "Nucleotide sequence of cloned cDNA for human sphingolipid activator  
RT protein 1 precursor.";  
RL Proc. Natl. Acad. Sci. U.S.A. 84:8652-8656 (1987).  
RN [6]  
RP PARTIAL SEQUENCE OF 60-142.  
RX MEDLINE=89240739; PubMed=2717620;  
RA Morimoto S., Martin B.M., Yamamoto Y., Kretz K.A., O'Brien J.S.,  
RA Kishimoto Y.;  
RT "Saposin A: second cerebroside activator protein.";  
RL Proc. Natl. Acad. Sci. U.S.A. 86:3389-3393 (1989).  
RN [7]  
RP SEQUENCE OF 195-263 FROM N.A.  
RX MEDLINE=86130593; PubMed=2869718;  
RA Dewji N.N., Wenger D.A., Fujibayashi S., Donoviel M., Esch F.,  
RA Hill F., O'Brien J.S.;  
RT "Molecular cloning of the sphingolipid activator protein-1 (SAP-1),  
RT the sulfate sulfate activator.";  
RL Biochem. Biophys. Res. Commun. 134:989-994 (1986).  
RN [8]  
RP SEQUENCE OF 195-274.  
RC TISSUE=Kidney;  
RX MEDLINE=91006165; PubMed=2209618;  
RA Furst W., Schubert J., Machleidt W., Meyer H.E., Sandhoff K.;  
RT "The complete amino-acid sequences of human ganglioside GM2 activator  
RT protein and cerebroside sulfate activator protein.";  
RL Eur. J. Biochem. 192:709-714 (1990).  
RN [9]  
RP SEQUENCE OF 195-274.  
RX MEDLINE=89207118; PubMed=3242555;  
RA Kleinschmidt T., Christomanou H., Braunitzer G.;  
RT "Complete amino-acid sequence of the naturally occurring A2 activator  
RT protein for enzymic sphingomyelin degradation: identity to the  
RT sulfate activator protein (SAP-1).";  
RL Biol. Chem. Hoppe-Seyler 369:1361-1365 (1988).  
RN [10]  
RP SEQUENCE OF 311-390.  
RX MEDLINE=88163077; PubMed=3442600;  
RA Kleinschmidt T., Christomanou H., Braunitzer G.;  
RT "Complete amino-acid sequence and carbohydrate content of the  
RT naturally occurring glucosylceramide activator protein (A1 activator)  
RT absent from a new human Gaucher disease variant.";  
RL Biol. Chem. Hoppe-Seyler 368:1571-1578 (1987).  
RN [11]  
RP SEQUENCE OF 407-484.  
RX MEDLINE=89000190; PubMed=3048308;  
RA Furst W., Machleidt W., Sandhoff K.;  
RT "The precursor of sulfate activator protein is processed to three  
RT different proteins.";  
RL Biol. Chem. Hoppe-Seyler 369:317-328 (1988).  
RN [12]  
RP PARTIAL SEQUENCE OF 405-484.  
RX MEDLINE=89025876; PubMed=2845979;  
RA Morimoto S., Martin B.M., Kishimoto Y., O'Brien J.S.;  
RT "Saposin D: a sphingomyelinase activator.";  
RL Biochem. Biophys. Res. Commun. 156:403-410 (1988).  
RN [13]  
RP SEQUENCE OF 17-26.  
RC TISSUE=Milk;  
RX MEDLINE=92068206; PubMed=1958198;  
RA Kondoh K., Hineno T., Sano A., Kakimoto Y.;  
RT "Isolation and characterization of prosaposin from human milk.";  
RL Biochem. Biophys. Res. Commun. 181:286-292 (1991).  
RN [14]  
RP PARTIAL SEQUENCE (SAPOSIN B), AND STRUCTURE OF CARBOHYDRATES.  
RC TISSUE=Urine;  
RX MEDLINE=20032116; PubMed=10562467;  
RA Fluhrty A.L., Lombardo C., Louis A., Stevens R.L., Whitelegge J.P.,  
RA Waring A.J., To T., Fluhrty C.B., Faull K.F.;  
RT "Preparation of the cerebroside sulfate activator (CSAct or saposin B)  
RT from human urine.";  
RL Mol. Genet. Metab. 68:391-403 (1999).  
RN [15]  
RP STRUCTURE OF CARBOHYDRATE ON ASN-215.  
RX MEDLINE=21110404; PubMed=11180632;  
RA Faull K.F., Johnson J., Kim M.J., To T., Whitelegge J.P.,  
RA Stevens R.L., Fluhrty C.B., Fluhrty A.L.;  
RT "Structure of the asparagine-linked sugar chains of porcine kidney and  
RT human urine cerebroside sulfate activator protein.";  
RL J. Mass Spectrom. 35:1416-1424 (2000).  
RN [16]  
RP SAPOSIN D DISULFIDE BONDS.  
RX MEDLINE=99337688; PubMed=10406958;  
RA Tatti M., Salvioli R., Ciaffoni F., Pucci P., Andolfo A.,  
RA Amoresano A., Vaccaro A.M.;  
RT "Structural and membrane-binding properties of saposin D.";  
RL Eur. J. Biochem. 263:486-494 (1999).  
RN [17]  
RP SAPOSIN B DISULFIDE BONDS.  
RX MEDLINE=22398398; PubMed=12510003;  
RA Ahn V.E., Faull K.F., Whitelegge J.P., Higginson J., Fluhrty A.L.,  
RA Prive G.G.;  
RT "Expression, purification, crystallization, and preliminary X-ray  
RT analysis of recombinant human saposin B.";  
RL Protein Expr. Purif. 27:186-193 (2003).  
RN [18]  
RP MASS SPECTROMETRY.  
RC TISSUE=Urine;  
RX MEDLINE=99441404; PubMed=10510427;  
RA Faull K.F., Whitelegge J.P., Higginson J., To T., Johnson J.,  
RA Krutshinsky A.N., Standing K.G., Waring A.J., Stevens R.L.,  
RA Fluhrty C.B., Fluhrty A.L.;  
RT "Cerebroside sulfate activator protein (Saposin B): chromatographic  
RT and electrospray mass spectrometric properties.";  
RL J. Mass Spectrom. 34:1040-1054 (1999).  
RN [19]  
RP X-RAY CRYSTALLOGRAPHY (2.2 ANGSTROMS) OF 195-273, AND MUTAGENESIS OF  
RP ILE-240.  
RX MEDLINE=22406333; PubMed=12518053;  
RA Ahn V.E., Faull K.F., Whitelegge J.P., Fluhrty A.L., Prive G.G.;  
RT "Crystal structure of saposin B reveals a dimeric shell for lipid  
RT binding.";  
RL Proc. Natl. Acad. Sci. U.S.A. 100:38-43 (2003).  
RN [20]  
RP REVIEW ON MLD VARIANTS.  
RX MEDLINE=95170731; PubMed=7866401;  
RA Gieselmann V., Zlotogora J., Harris A., Wenger D.A., Morris C.P.;  
RT "Molecular genetics of metachromatic leukodystrophy.";  
RL Hum. Mutat. 4:233-242 (1994).  
RN [21]  
RP VARIANT MLD ILE-217.  
RX MEDLINE=90147748; PubMed=2302219;  
RA Rafi M.A., Zhang X.-L., Degala G., Wenger D.A.;  
RT "Detection of a point mutation in sphingolipid activator protein-1  
RT mRNA in patients with a variant form of metachromatic  
RT leukodystrophy.";  
RL Biochem. Biophys. Res. Commun. 166:1017-1023 (1990).  
RN [22]  
RP SEQUENCE FROM N.A., AND VARIANT MLD ILE-217.  
RX MEDLINE=90207231; PubMed=2320574;

```
Query Match      100.0%; Score 524; DB 1; Length 524;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 524; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MYALFILASLLGAALAGPVLGKCTRGSAVWCQVKTASDCGAKVHCLQTVWKNKPTVKS 60
Db 1 MYALFILASLLGAALAGPVLGKCTRGSAVWCQVKTASDCGAKVHCLQTVWKNKPTVKS 60
Qy 61 LPDCICKDVVTAAGDMLKDNATEEILVYLEKTCMDLPPKPNMSASCKEIVDSYLPVILDI 120
Db 61 LPDCICKDVVTAAGDMLKDNATEEILVYLEKTCMDLPPKPNMSASCKEIVDSYLPVILDI 120
Qy 121 IKGMSRPGVCSALNLCESLQKHLAELNHQKLESNKIPELDMTEVVAPFMANIPLLLY 180
Db 121 IKGMSRPGVCSALNLCESLQKHLAELNHQKLESNKIPELDMTEVVAPFMANIPLLLY 180
Qy 181 PQDGRSKPQKONGDVQCDCIQMTDIOTAVRTNSTFVQALVEHVKEECDRLPGMADI 240
Db 181 PQDGRSKPQKONGDVQCDCIQMTDIOTAVRTNSTFVQALVEHVKEECDRLPGMADI 240
Qy 241 CKNYISYSEIAIQMMHMPQKEICALVGFCDVEKEMPMOTLVPKAVSKNVPALVELVE 300
Db 241 CKNYISYSEIAIQMMHMPQKEICALVGFCDVEKEMPMOTLVPKAVSKNVPALVELVE 300
Qy 301 PIKHEVPAKSDVYCEVCEFLVKEVTKLIDNNKTEKEILDADFQKMSKLPKSLSEECQEV 360
Db 301 PIKHEVPAKSDVYCEVCEFLVKEVTKLIDNNKTEKEILDADFQKMSKLPKSLSEECQEV 360
Qy 361 VDTYSSILSILLBEVSPVLCVSMHLCSGTRLPALTTHVTPQKGGFCEVCKKLGVYLD 420
Db 361 VDTYSSILSILLBEVSPVLCVSMHLCSGTRLPALTTHVTPQKGGFCEVCKKLGVYLD 420
Qy 421 RNLKNTKTEILAALEKGCSEFLPDYKQKCDQFVAEYEPVLIELVEVMDPSVCLKIG 480
Db 421 RNLKNTKTEILAALEKGCSEFLPDYKQKCDQFVAEYEPVLIELVEVMDPSVCLKIG 480
Qy 481 APCSAHPLLGTEKXICWGPSYWCQNTETAACQNAVEHCKRHVWN 524
Db 481 APCSAHPLLGTEKXICWGPSYWCQNTETAACQNAVEHCKRHVWN 524

RESULT 2
SAP_BOVIN
ID - SAP BOVIN STANDARD; PRT; 525 AA.
AC P26779; Q9N2G4;
DT 01-AUG-1992 (Rel. 23, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DE Proactivator polypeptide precursor [Contains: Saposin A (Protein A);
DE Saposin B (Sphingolipid activator protein 1) (SAP-1) (Cerebroside
DE sulfate activator) (CSact) (Dispersin) (Sulfatide/GMI activator);
DE Saposin C (Co-beta-glucosidase) (Al activator) (Glucosylceramidase
DE activator) (Sphingolipid activator protein 2) (SAP-2); Saposin D
DE (Protein C) (Component C)].
GN PSAP.
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovidae; Bovinae; Bos.
OX NCBI_TaxID=9913;
RN [1]
RP SEQUENCE FROM N.A., AND VARIANTS.
RC TISSUE=Mammary gland;
RA Azuma N., Yoshida K.;
RT "RT-PCR cloning of bovine proasaposin.";
RL Submitted (JAN-2000) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE OF 312-391.
RC TISSUE=Spleen;
RA MEDLINE=92207994; PubMed=1554743;
RA Sano A., Mizuno T., Kondoh K., Hineno T., Ueno S.-I., Kakimoto Y.,
RA Morita N.;
```

```
RT "Saposin-C from bovine spleen; complete amino acid sequence and
RT relation between the structure and its biological activity.";
RL Biochim. Biophys. Acta 1120:75-80(1992).
CC -i- FUNCTION: The lysosomal degradation of sphingolipids takes place
CC by the sequential action of specific hydrolases. Some of these
CC enzymes require specific low-molecular mass, non-enzymic proteins:
CC the sphingolipids activator proteins (coproteins) (By similarity).
CC -i- FUNCTION: Saposin A and saposin C stimulate the hydrolysis of
CC galactosylceramide by beta-glucosylceramidase (EC 3.2.1.45) and
CC galactosylceramide by beta-galactosylceramidase (EC 3.2.1.46).
CC Saposin C apparently acts by combining with the enzyme and acidic
CC lipid to form an activated complex, rather than by solubilizing
CC the substrate.
CC -i- FUNCTION: Saposin B stimulates the hydrolysis of galacto-
CC cerebroside sulfate by arylsulfatase A (EC 3.1.1.6.8), GMI
CC gangliosides by beta-galactosidase (EC 3.2.1.23) and
CC globotriaosylceramide by alpha-galactosidase A (EC 3.2.1.22).
CC Saposin B forms a solubilizing complex with the substrates of the
CC sphingolipid hydrolases (By similarity).
CC -i- FUNCTION: Saposin D is a specific sphingomyelin phosphodiesterase
CC activator (EC 3.1.4.12) (By similarity).
CC -i- SUBUNIT: Saposin B is a homodimer (By similarity).
CC -i- SUBCELLULAR LOCATION: Lysosomal.
CC -i- PTM: This precursor is proteolytically processed to 4 small
CC peptides, which are similar to each other and are sphingolipid
CC hydrolase activator proteins (By similarity).
CC -i- SIMILARITY: Contains 2 saposin A-type domains.
CC -i- SIMILARITY: Contains 4 saposin B-type domains.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
EMBL; AB036791; BAA95677.1; -.
DR InterPro; IPR003119; SapA.
DR InterPro; IPR007856; SapB_1.
DR InterPro; IPR008138; SapB_2.
DR InterPro; IPR008140; SapB_sub.
DR InterPro; IPR008373; Saposin.
DR InterPro; IPR008139; SaposinB.
DR Pfam; PF02199; SAPA; 2.
DR Pfam; PF05184; SapB_1; 4.
DR Pfam; PF03489; SapB_2; 4.
DR PRINTS; PR01797; SAPOSIN.
DR ProDom; PD001732; SapB_sub; 3.
DR SMART; SM00162; SAPA; 2.
DR SMART; SM00118; SAPB; 4.
KW Signal; Glycoprotein; Lysosome; Sphingolipid metabolism; Repeat.
FT SIGNAL 1 16
FT PROPEP 17 58
FT CHAIN 60 142 SAPOSIN A.
FT PROPEP 144 195
FT CHAIN 196 275 SAPOSIN B.
FT PROPEP 277 310
FT CHAIN 312 392 SAPOSIN C.
FT PROPEP 393 404
FT CHAIN 406 487 SAPOSIN D.
FT PROPEP 489 525
FT DOMAIN 21 54
FT DOMAIN 59 142 SAPOSIN-LIKE TYPE A 1.
FT DOMAIN 194 276 SAPOSIN-LIKE TYPE B 1.
FT DOMAIN 312 393 SAPOSIN-LIKE TYPE B 2.
FT DOMAIN 406 487 SAPOSIN-LIKE TYPE B 3.
FT DOMAIN 492 525 SAPOSIN-LIKE TYPE B 4.
FT DISULFID 63 138 SAPOSIN-LIKE TYPE A 2.
FT DISULFID 66 132 BY SIMILARITY.
FT DISULFID 94 106 BY SIMILARITY.
FT DISULFID 199 272 BY SIMILARITY.
FT DISULFID 202 266 BY SIMILARITY.
```

```

FT DISULFID 231 242 BY SIMILARITY.
FT DISULFID 316 389 BY SIMILARITY.
FT DISULFID 319 383 BY SIMILARITY.
FT DISULFID 347 358 BY SIMILARITY.
FT DISULFID 410 483 BY SIMILARITY.
FT DISULFID 413 477 BY SIMILARITY.
FT DISULFID 441 452 BY SIMILARITY.
FT CARBOHYD 80 80 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 101 101 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 216 236 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 333 333 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 427 427 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT VARIANT 127 127 H -> R.
FT VARIANT 260 263 MOPK -> IRIK.
FT CONFLICT 317 317 E -> Q (IN REF. 2).
FT CONFLICT 367 367 R -> S (IN REF. 1).
SQ SEQUENCE 525 AA; 58120 MW; 293AFC0F9C4FA99 CRC64;

Query Match 5.7%; Score 30; DB 1; Length 525;
Best Local Similarity 100.0%; Pred. No. 8.7e-22;
Matches 30; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 128 PGEVCSALNLCESLOKHLAELNHQKLESN 157
DB 128 PGEVCSALNLCESLOKHLAELNHQKLESN 157

RESULT 3
SAP_PIG STANDARD; PRT; 80 AA.
AC P81405;
DT 15-DEC-1998 (Rel. 37, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Saposin B (Cerebroside sulfate activator) (CS-ACT) (Non-specific
DE activator) (Sphingolipid activator protein 1) (SAP-1).
OS Sus scrofa (Fig.).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
OX NCBI_TaxID=9823;
RN [1]
RP SEQUENCE OF 1-79.
RC TISSUE=Kidney;
RX MEDLINE=92229506; PubMed=8471613;
RA Stevens R.L., Faull K.F., Conklin K.A., Green B.N., Fluharty A.L.;
RT "Porcine cerebroside sulfate activator; further structural
RT characterization and disulfide identification.";
RL Biochemistry 32:4051-4059(1993).
RN [2]
RP SEQUENCE OF 1-64.
RC TISSUE=Kidney;
RX MEDLINE=9222651; PubMed=1562358;
RA Fluharty A.L., Katona Z., Meek W.E., Frei K., Fowler A.V.;
RT "The cerebroside sulfate activator from pig kidney: purification and
RT molecular structure.";
RL Biochem. Med. Metab. Biol. 47:66-85(1992).
RN [3]
RP STRUCTURE OF CARBOHYDRATE ON ASN-21.
RX MEDLINE=21110404; PubMed=11180632;
RA Faull K.F., Johnson J., Kim M.J., To T., Whitelegge J.P.,
RA Stevens R.L., Fluharty C.B., Fluharty A.L.;
RT "Structure of the asparagine-linked sugar chains of porcine kidney and
RT human urine cerebroside sulfate activator protein.";
RL J. Mass Spectrom. 35:1416-1424(2000).
RN [4]
RP MASS SPECTROMETRY.
RC TISSUE=Kidney;
RX MEDLINE=99441404; PubMed=10510427;
RA Faull K.F., Whitelegge J.P., Higginson J., To T., Johnson J.,
RA Krutchinsky A.N., Standing K.G., Waring A.J., Stevens R.L.,
RA Fluharty C.B., Fluharty A.L.;
RT "Cerebroside sulfate activator protein (Saposin B): chromatographic
RT and electrospray mass spectrometric properties.";

```

```

RL J. Mass Spectrom. 34:1040-1054(1999).
CC -1- FUNCTION: Saposin B stimulates the hydrolysis of galacto-
CC cerebroside sulfate by arylsulfatase A (EC 3.1.6.8), GM1
CC gangliosides by beta-galactosidase (EC 3.2.1.23) and
CC globotriaosylceramide by alpha-galactosidase A (EC 3.2.1.22).
CC Saposin B forms a solubilizing complex with the substrates of the
CC sphingolipid hydrolases.
CC -1- SUBUNIT: Saposin B is a homodimer (By similarity).
CC -1- PTM: The one residue extended Saposin B-Val is only found in a
CC minority of the chains.
CC -1- SIMILARITY: Contains 1 saposin B-type domain.
DR GlycoSuiteDB; P81405; -.
DR InterPro; IPR007856; SapB 1.
DR InterPro; IPR008138; SapB 2.
DR InterPro; IPR008373; Saposin.
DR InterPro; IPR008139; SaposinB.
DR Pfam; PF05184; SapB_1; 1.
DR Pfam; PF03489; SapB_2; 1.
DR PRINTS; PR01797; SAPOSIN.
DR SMART; SM00118; SAPB; 1.
KW Glycoprotein; Sphingolipid metabolism.
FT CHAIN 1 79
FT CHAIN 1 80 SAPOSIN B.
FT DOMAIN 1 80 SAPOSIN B-VAL.
FT DISULFID 4 77 SAPOSIN-LIKE TYPE B.
FT DISULFID 7 71
FT DISULFID 36 47
FT CARBOHYD 21 21 N-LINKED (GLCNAC. . .) (COMPLEX).
FT SEQUENCE 80 AA; 8949 MW; EF7BA249863E789C CRC64;
SQ SEQUENCE 80 AA; 8949 MW; EF7BA249863E789C CRC64;

Query Match 4.2%; Score 22; DB 1; Length 80;
Best Local Similarity 100.0%; Pred. No. 1.8e-14;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 241 CKNVISQYSEIAIQMMHMQPK 262
DB 47 CKNVISQYSEIAIQMMHMQPK 69

RESULT 4
SAP_RAT STANDARD; PRT; 554 AA.
AC P10560; Q62841; Q64190;
DT 01-JUL-1989 (Rel. 11, Created)
DT 01-JUL-1989 (Rel. 11, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE Sulfated glycoprotein 1 precursor (SGP-1) (Prosaposin).
GN PSAP OR SGP1.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
RC TISSUE=Sertoli cells;
RX MEDLINE=8900647; PubMed=3048385;
RA Collard M.W., Sylvester S.R., Tsuruta J.K., Griswold M.D.;
RT "Biosynthesis and molecular cloning of sulfated glycoprotein 1
RT secreted by rat Sertoli cells; sequence similarity with the
RT 70-kilodalton precursor to sulfatide/GM1 activator.";
RL Biochemistry 27:4557-4564(1988).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Testis;
RX MEDLINE=96128541; PubMed=8573994;
RA Morales C.R., El-Alfy M., Zhao Q., Igdown S.A.;
RT "Molecular role of sulfated glycoprotein-1 (SGP-1/prosaposin) in
RT Sertoli cells.";
RL Histol. Histopathol. 10:1023-1034(1995).
RN [3]
RP SEQUENCE FROM N.A.
RC TISSUE=Testis;

```

```
RX MEDLINE=96175245; PubMed=8601692;
RA Morales C.R., El-Alfy M., Zhao Q., Igdoura S.A.;
RT "Expression and tissue distribution of rat sulfated glycoprotein-1
  (prosaposin).";
RL J. Histochem. Cytochem. 44:327-337(1996).
CC -!- SUBCELLULAR LOCATION: Extracellular.
CC -!- SIMILARITY: Contains 2 saposin A-type domains.
CC -!- SIMILARITY: Contains 4 saposin B-type domains.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; M19936; AAA42136.1; -
DR EMBL; S81353; AAB36042.2; -
DR EMBL; S81373; AAB36233.2; -
DR PIR; A28716; A28716.
DR InterPro; IPR003119; Sapa.
DR InterPro; IPR007856; SapB_1.
DR InterPro; IPR008138; SapB_2.
DR InterPro; IPR008140; SapB_sub.
DR InterPro; IPR008373; Saposin.
DR InterPro; IPR008139; SaposinB.
DR Pfam; PF02199; SAPA; 2.
DR Pfam; PF05184; SapB_1; 3.
DR Pfam; PF02199; SAPA; 2.
DR Pfam; PF03489; SapB_2; 4.
DR PRINTS; PR01797; SAPOSIN.
DR ProDom; PD001732; SapB_sub; 3.
DR SMART; SM00162; SAPA; 2.
DR SMART; SM00118; SApB; 4.
DR Signal; Glycoprotein; Repeat.
FT SIGNAL 1 16
FT CHAIN 17 554
FT DOMAIN 21 54
FT DOMAIN 61 138
FT DOMAIN 193 274
FT DOMAIN 310 391
FT DOMAIN 435 516
FT DOMAIN 521 554
FT DISULFID 63 138
FT DISULFID 66 132
FT DISULFID 94 106
FT DISULFID 197 270
FT DISULFID 200 264
FT DISULFID 229 240
FT DISULFID 314 387
FT DISULFID 317 381
FT DISULFID 345 356
FT DISULFID 439 512
FT DISULFID 442 506
FT DISULFID 470 481
FT CARBOHYD 80 80
FT CARBOHYD 214 214
FT CARBOHYD 331 331
FT CARBOHYD 456 456
FT CARBOHYD 115 115
FT CONFLICT 299 299
FT CONFLICT 462 462
FT CONFLICT 527 527
FT CONFLICT 536 536
SQ SEQUENCE 554 AA; 61123 MW; DFE3F3A3A0520C6B CRC64;
Query Match 4.2%; Score 22; DB 1; Length 554;
Best Local Similarity 100.0%; Pred. No. 9.6e-14;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 431 EILAALEKGCFLPDPYQKQCD 452
Db 461 EILAALEKGCFLPDPYQKQCD 482
```

```
RESULT 5
ID SAP MOUSE STANDARD; PRT; 557 AA.
AC Q61207; Q60861; Q64006; Q64219;
DT 15-DEC-1998 (Rel. 37, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE Sulfated glycoprotein 1 precursor (SGP-1) (Prosaposin).
GN PSAP OR SGPI
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Liver;
RX MEDLINE=92272718; PubMed=1590788;
RA Teuda M., Sakiyama T., Endo H., Kitagawa T.;
RT "The primary structure of mouse saposin.";
RL Biochem. Biophys. Res. Commun. 184:1266-1272(1992).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=94272317; PubMed=8003952;
RA Sprecher-Levy H., Orr-Urtreger A., Lonai P., Horowitz M.;
RT "Murine prosaposin: expression in the reproductive system of a gene
  implicated in human genetic disease.";
RL Cell. Mol. Biol. 40:233-233(1994).
RN [3]
RP SEQUENCE FROM N.A.
RX MEDLINE=96084310; PubMed=8565332;
RA Cao Q.P., Crain W.R.;
RT "Expression of SGP-1 mRNA in preimplantation mouse embryos.";
RL Dev. Genet. 17:263-271(1995).
RN [4]
RP SEQUENCE FROM N.A.
RC STRAIN=BALB/c; TISSUE=Liver;
RA Zhao Q.Q., Hay N.N., Morales C.R.;
RL Submitted (MAY-1996) to the EMBL/GenBank/DBJ databases.
CC -!- SUBCELLULAR LOCATION: Extracellular.
CC -!- SIMILARITY: Contains 2 saposin A-type domains.
CC -!- SIMILARITY: Contains 4 saposin B-type domains.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; S36200; AAB22175.1; -
DR EMBL; S71616; AAB31059.1; -
DR EMBL; U27340; AAA92567.1; -
DR PIR; JH0604; JH0604.
DR MGD; MGI:97783; Psap.
DR InterPro; IPR003119; Sapa.
DR InterPro; IPR007856; SapB_1.
DR InterPro; IPR008138; SapB_2.
DR InterPro; IPR008140; SapB_sub.
DR InterPro; IPR008373; Saposin.
DR InterPro; IPR008139; SaposinB.
DR Pfam; PF02199; SAPA; 2.
DR Pfam; PF05184; SapB_1; 4.
DR Pfam; PF03489; SapB_2; 4.
DR PRINTS; PR01797; SAPOSIN.
DR ProDom; PD001732; SapB_sub; 3.
DR SMART; SM00162; SAPA; 2.
DR SMART; SM00118; SApB; 4.
DR Signal; Glycoprotein; Repeat.
FT SIGNAL 1 16
BY SIMILARITY.
```

FT CHAIN 17 557  
 FT DOMAIN 21 54  
 FT DOMAIN 59 142  
 FT DOMAIN 193 277  
 FT DOMAIN 313 394  
 FT DOMAIN 438 519  
 FT DOMAIN 524 557  
 FT DISULFID 63 138  
 FT DISULFID 66 132  
 FT DISULFID 94 106  
 FT DISULFID 197 273  
 FT DISULFID 200 267  
 FT DISULFID 229 240  
 FT DISULFID 317 390  
 FT DISULFID 320 384  
 FT DISULFID 348 359  
 FT DISULFID 442 515  
 FT DISULFID 445 509  
 FT DISULFID 473 484  
 FT CARBOHYD 80 80  
 FT CARBOHYD 214 214  
 FT CARBOHYD 334 334  
 FT CARBOHYD 459 459  
 FT CONFLICT 83 83  
 FT CONFLICT 158 158  
 FT CONFLICT 160 160  
 FT CONFLICT 171 172  
 FT CONFLICT 244 244  
 FT CONFLICT 254 254  
 FT CONFLICT 255 255  
 FT CONFLICT 260 262  
 FT CONFLICT 307 307  
 FT CONFLICT 322 322  
 FT CONFLICT 349 350  
 FT CONFLICT 367 367  
 FT CONFLICT 370 370  
 FT CONFLICT 373 373  
 FT CONFLICT 391 391  
 FT CONFLICT 393 393  
 FT CONFLICT 406 406  
 FT CONFLICT 430 430  
 FT CONFLICT 445 445  
 FT CONFLICT 448 448  
 SQ SEQUENCE 557 AA; 61422 MW; 134593E20499535E CRC64;

Query Match 4.2%; Score 22; DB 1; Length 557;  
 Best Local Similarity 100.0%; Pred. No. 9.7e-14;  
 Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 431 EILAALEKGCFLDPDPYQKQCD 452  
 Db 464 EILAALEKGCFLDPDPYQKQCD 485

RESULT 6  
 ID - SAP CHICK STANDARD; PRT; 518 AA.  
 AC O13035;  
 DT 16-OCT-2001 (Rel. 40, Created)  
 DT 16-OCT-2001 (Rel. 40, Last sequence update)  
 DT 28-FEB-2003 (Rel. 41, Last annotation update)  
 DE Proactivator polypeptide precursor [Contains: Saposin A; Saposin B; Saposin C; Saposin D].  
 GN PSAP.  
 OS Gallus gallus (Chicken).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;  
 OC Gallus.  
 OX NCBI\_TaxID=9031;  
 RN [1]  
 RP SEQUENCE FROM N.A., AND SEQUENCE OF 194-203.  
 RC TISSUE=Brain, and Liver;  
 RX MEDLINE=98129745; PubMed=9461526;

RA Azuma N., Seo H.-C., Lie O., Fu Q., Gould R.M., Hiraiwa M., Burt D.W.,  
 RA Paton I.R., Morrice D.R., O'Brien J.S., Kishimoto Y.,  
 RT "Cloning, expression and map assignment of chicken prosaposin.";  
 RL Biochem. J. 330:321-327(1998).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RA Altman N., Horowitz M.;  
 RL Submitted (NOV-1998) to the EMBL/GenBank/DBJ databases.  
 CC -!- FUNCTION: The lysosomal degradation of sphingolipids takes place  
 CC by the sequential action of specific hydrolases. Some of these  
 CC enzymes require specific low-molecular mass, non-enzymic proteins:  
 CC the sphingolipids activator proteins (coproteins) (By similarity).  
 CC -!- FUNCTION: Saposin A and saposin C stimulate the hydrolysis of  
 CC glucosylceramide by beta-glucosylceramidase (EC 3.2.1.45) and  
 CC galactosylceramide by beta-galactosylceramidase (EC 3.2.1.46).  
 CC Saposin C apparently acts by combining with the enzyme and acidic  
 CC lipid to form an activated complex, rather than by solubilizing  
 CC the substrate (By similarity).  
 CC -!- FUNCTION: Saposin B stimulates the hydrolysis of galacto-  
 CC cerebroside sulfate by arylsulfatase A (EC 3.1.6.8), GM1  
 CC gangliosides by beta-galactosidase (EC 3.2.1.23) and  
 CC globotriaosylceramide by alpha-galactosidase A (EC 3.2.1.22).  
 CC Saposin B forms a solubilizing complex with the substrates of the  
 CC sphingolipid hydrolases (By similarity).  
 CC -!- FUNCTION: Saposin D is a specific sphingomyelin phosphodiesterase  
 CC activator (EC 3.1.4.12) (By similarity).  
 CC -!- SUBUNIT: Saposin B is a homodimer (By similarity).  
 CC -!- SUBCELLULAR LOCATION: Lysosomal (By similarity).  
 CC -!- PTM: This precursor is proteolytically processed to 4 small  
 CC peptides, which are similar to each other and are sphingolipid  
 CC hydrolase activator proteins (By similarity).  
 CC -!- SIMILARITY: Contains 2 saposin A-type domains.  
 CC -!- SIMILARITY: Contains 4 saposin B-type domains.

This SWISS-PROT entry is copyright. It is produced through a collaboration  
 between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 the European Bioinformatics Institute. There are no restrictions on its  
 use by non-profit institutions as long as its content is in no way  
 modified and this statement is not removed. Usage by and for commercial  
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

EMBL; AB003471; BAA19914.1; -  
 EMBL; AF108656; AAF05899.1; -  
 DR InterPro; IPR003119; Sapa.  
 DR InterPro; IPR007856; SapB\_1.  
 DR InterPro; IPR008138; SapB\_2.  
 DR InterPro; IPR008140; SapB\_sub.  
 DR InterPro; IPR008373; Saposin.  
 DR InterPro; IPR008139; SaposinB.  
 DR Pfam; PF02199; SAPA; 2.  
 DR Pfam; PF05184; SapB\_1; 4.  
 DR Pfam; PF03489; SapB\_2; 4.  
 DR PRINTS; PR01797; SAPOSIN.  
 DR ProDom; PD001732; SapB\_sub; 3.  
 KW Signal; Glycoprotein; Lysosome; Sphingolipid metabolism; Repeat;  
 KW GM2-gangliosidosis.  
 FT SIGNAL 1 17 POTENTIAL.  
 FT PROPEP 18 60  
 FT CHAIN 61 143 SAPOSIN A.  
 FT PROPEP 145 193  
 FT CHAIN 194 276 SAPOSIN B.  
 FT PROPEP 278 305  
 FT CHAIN 307 387 SAPOSIN C.  
 FT PROPEP 389 398  
 FT CHAIN 399 480 SAPOSIN D.  
 FT PROPEP 482 518  
 FT DOMAIN 22 55  
 FT DOMAIN 60 143 SAPOSIN-LIKE TYPE A 1.  
 FT DOMAIN 193 277 SAPOSIN-LIKE TYPE B 1.  
 FT DOMAIN 307 388 SAPOSIN-LIKE TYPE B 2.  
 FT DOMAIN 399 480 SAPOSIN-LIKE TYPE B 3.  
 FT DOMAIN 485 518 SAPOSIN-LIKE TYPE A 2.

```
FT DISULFID 64 139 BY SIMILARITY.
FT DISULFID 67 133 BY SIMILARITY.
FT DISULFID 95 107 BY SIMILARITY.
FT DISULFID 197 273 BY SIMILARITY.
FT DISULFID 200 267 BY SIMILARITY.
FT DISULFID 229 240 BY SIMILARITY.
FT DISULFID 311 384 BY SIMILARITY.
FT DISULFID 314 378 BY SIMILARITY.
FT DISULFID 342 353 BY SIMILARITY.
FT DISULFID 403 476 BY SIMILARITY.
FT DISULFID 406 470 BY SIMILARITY.
FT DISULFID 434 445 BY SIMILARITY.
FT CARBOHYD 81 81 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 214 214 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 328 328 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 420 420 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CONFLICT 94 94 R -> T (IN REF. 2).
FT CONFLICT 486 486 E -> D (IN REF. 2).
SQ SEQUENCE 518 AA; 57601 MW; B803000E891C3963 CRC64;

Query Match 2.3%; Score 12; DB 1; Length 518;
Best Local Similarity 100.0%; Pred. No. 0.00096;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 106 CKEIVDSYLPVI 117
Db 107 CKEIVDSYLPVI 118

RESULT 7
SAP_CAVPO STANDARD; PRT; 81 AA.
ID -SAP_CAVPO
AC P20097;
DT 01-FEB-1991 (Rel. 17, Created)
DT 01-FEB-1991 (Rel. 17, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Saposin C (CO-beta-glucosidase) (Glucosylceramide activator)
DE (Sphingolipid activator protein 2) (SAP-2).
OS Cavia porcellus (Guinea pig).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Hystriognathi; Caviidae; Cavia.
OX NCBI_TaxID=10141;
RN [1]
RP SEQUENCE.
RC TISSUE=Liver;
RX MEDLINE=89066787; PubMed=3198642;
RA Sano A., Radin N.S., Johnson L.L., Tarr G.E.;
RT "The activator protein for glucosylceramide beta-glucosidase from
RT guinea pig liver. Improved isolation method and complete amino acid
RT sequence."
RL J. Biol. Chem. 263:19597-19601(1988).
CC -1- FUNCTION: Saposin A and saposin C stimulate the hydrolysis of
CC glucosylceramide by beta-glucosylceramidase (EC 3.2.1.45) and
CC galactosylceramide by beta-galactosylceramidase (EC 3.2.1.46).
CC Saposin C apparently acts by combining with the enzyme and acidic
CC lipid to form an activated complex, rather than by solubilizing
CC the substrate.
CC -1- SIMILARITY: Contains 1 saposin B-type domain.
DR PIR: A32026; A32026.
DR HSP; P42210; IQDM.
DR InterPro; IPR007856; SapB_1.
DR InterPro; IPR008138; SapB_2.
DR InterPro; IPR008140; SapB_sub.
DR InterPro; IPR008173; Saposin.
DR InterPro; IPR008139; SaposinB.
DR Pfam; PF05184; SapB_1; 1.
DR Pfam; PF03489; SapB_2; 1.
DR PRINTS; PR01797; SAPOSIN.
DR ProDom; PD001732; SapB_sub; 1.
DR SMART; SM00118; SAPB; 1.
KW Glycoprotein; Sphingolipid metabolism.
FT DOMAIN 1 81 SAPOSIN-LIKE TYPE B.
FT CARBOHYD 22 22 N-LINKED (GLCNAC. . .).
```

```
FT DISULFID 5 78 BY SIMILARITY.
FT DISULFID 8 72 BY SIMILARITY.
FT DISULFID 36 47 BY SIMILARITY.
SQ SEQUENCE 81 AA; 8852 MW; E564CE1F0A292596 CRC64;

Query Match 1.7%; Score 9; DB 1; Length 81;
Best Local Similarity 100.0%; Pred. No. 0.19;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 357 CQEVDTYG 365
Db 47 CQEVDTYG 55

RESULT 8
FIMI_ECOLI STANDARD; PRT; 179 AA.
ID -FIMI_ECOLI
AC P39264;
DT 01-FEB-1995 (Rel. 31, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Fimbrin-like protein fimi precursor.
DE Fimbrin-like protein fimi precursor.
GN FIMI OR B4315.
OS Escherichia coli.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Escherichia.
OX NCBI_TaxID=562;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=K12 / MG1655;
RX MEDLINE=95334362; PubMed=7610040;
RA Burland V.D., Plunkett G. III, Sofia H.J., Daniels D.L.,
RA Blattner F.R.;
RT "Analysis of the Escherichia coli genome VI: DNA sequence of the
RT region from 92.8 through 100 minutes."
RL Nucleic Acids Res. 23:2105-2119(1995).
RN [2]
RP PRELIMINARY SEQUENCE OF 1-80 FROM N.A.
RX MEDLINE=84285425; PubMed=6147250;
RA Klemm P.;
RT "The fima gene encoding the type-1 fimbrial subunit of Escherichia
RT coli. Nucleotide sequence and primary structure of the protein."
RL Eur. J. Biochem. 143:395-399(1984).
CC -1- SIMILARITY: BELONGS TO THE FIMA/PAPA FAMILY OF FIMBRIA PROTEINS.
CC -1- CAUTION: Ref.2 sequence differs from that shown due to
CC frameshifts.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; U14003; AAA97211.1; ALT INIT.
DR EMBL; AE000502; AAC77271.1; ALT INIT.
DR EMBL; X00981; -: NOT ANNOTATED_CDS.
DR EcoGene; EGI1974; fimi.
DR InterPro; IPR008966; Adhes_bact.
DR InterPro; IPR000259; Fimbrial.
DR Pfam; PF00419; Fimbrial; 1.
DR Pfam; Signal; Complete proteome.
FT SIGNAL 1 19 POTENTIAL.
FT CHAIN 20 179 FIMBRIN-LIKE PROTEIN FIMI.
FT DISULFID 43 83 BY SIMILARITY.
SQ SEQUENCE 179 AA; 19373 MW; ECE4FA7D24B85E25 CRC64;

Query Match 1.5%; Score 8; DB 1; Length 179;
Best Local Similarity 100.0%; Pred. No. 3.9;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 4 LFLASLL 11
```



```

Db 6 LFLASL 13
RESULT 9
HS23 DROME STANDARD; PRT; 186 AA.
AC P02516; Q9V5X5;
DT 21-JUL-1986 (Rel. 01, Created)
DT 01-NOV-1990 (Rel. 16, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Heat shock protein 23.
GN HSP23 OR CG4463.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
[1]
RN SEQUENCE FROM N.A.
RX MEDLINE=83189140; PubMed=6302284;
RA Southgate R., Ayme A., Voellmy R.;
RT "Nucleotide sequence analysis of the Drosophila small heat shock gene
RL cluster at locus 67B.";
RU J. Mol. Biol. 165:35-57(1983).
[2]
RN SEQUENCE FROM N.A.
RX MEDLINE=82248004; PubMed=6285380;
RA Ingolia T.D., Craig E.A.;
RT "Four small Drosophila heat shock proteins are related to each other
RL and to mammalian alpha-crystallin.";
RU Proc. Natl. Acad. Sci. U.S.A. 79:2360-2364(1982).
[3]
RN SEQUENCE FROM N.A.
RX STRAIN=Berkely;
RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Bernier B.P., Shandari D., Bolshakov S.,
RA Borokova D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,
RA Burtis K.C., Bueam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA de Pablo B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
RA Foaier C., Gabriellian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D.A., Heiman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Laoko P., Lei Y., Levitsky A.A., Li J.H., Li Z., Liang Y., Lin X.,
RA Markulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacle J.M.,
RA Palazolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Swirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Wassarman D.A., Weinstein G.M., Weissbach J.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong J., Zhou X., Zhou S., Zhu S., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT "The genome sequence of Drosophila melanogaster.";

```

---

```

RL Science 287:2185-2195 (2000).
RN [4]
RP SEQUENCE FROM N.A.
RC STRAIN=Berkely; TISSUE=Embryo;
RX MEDLINE=22426066; PubMed=12537569;
RA Stapleton M., Carlson J.W., Brokstein P., Yu C., Champe M.,
RA George R.A., Guarin H., Krommiller B., Pacle J.M., Park S., Wan K.H.,
RA Rubin G.M., Celniker S.E.;
RT "A Drosophila full-length cDNA resource.";
RL Genome Biol. 3:RESEARCH0080.1-RESEARCH0080.8(2002).
CC -i- SIMILARITY: Belongs to the small heat shock protein (HSP20)
CC family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; J01100; AAA28637.1; -
CC EMBL; V00210; CAA23494.1; -
CC EMBL; X03889; CAA27525.1; -
CC EMBL; A2003552; AAF50286.1; -
CC EMBL; AY061081; AAL28629.1; -
CC FIR; B20647; E20647.
CC Flybase; FBgn0001224; Hsp23.
CC InterPro; IPR001436; Crystallin_alpha.
CC InterPro; IPR002068; Hsp20.
CC InterPro; IPR008978; HSP20_chap.
CC Pfam; PF00011; HSP20; 1.
CC PRINTS; PR00299; ACRYSTALLIN.
CC PROSITE; PS01031; HSP20; 1.
KW Heat shock; Multigene family.
FT CONFLICT 31 33 QRN -> RRI (IN REF. 2).
FT CONFLICT 81 81 K -> E (IN REF. 2).
FT CONFLICT 88 88 K -> G (IN REF. 2).
FT CONFLICT 95 95 L -> V (IN REF. 2).
FT CONFLICT 173 173 N -> S (IN REF. 2).
FT CONFLICT 182 182 N -> G (IN REF. 2).
SQ SEQUENCE 186 AA; 20629 MW; 3C6EDCDAL19BSA0 CRC64;
Query Match 1.5%; Score 8; DB 1; Length 186;
Best Local Similarity 100.0%; Pred. No. 4;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 172 MANIPLLL 179
DB 1 MANIPLLL 8
RESULT 10
FMFX BORPE STANDARD; PRT; 201 AA.
AC P09808;
DT 01-MAR-1989 (Rel. 10, Created)
DT 10-OCT-2003 (Rel. 42, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE Fimbrial protein fimX precursor (Pilin).
GN FIMX OR BF2674.
OS Bordetella pertussis.
OC Bacteria; Proteobacteria; Betaproteobacteria; Burkholderiales;
OC Alcaligenaceae; Bordetella.
OX NCBI_TaxID=520;
[1]
RN SEQUENCE FROM N.A.
RC STRAIN=SAI;
RA MEDLINE=89013896; PubMed=2902506;
RA Pedroni P., Riboli B., de Ferra F., Grandi G., Toma S., Arico B.,
RA Rappuoli R.;
RT "Cloning of a novel pilin-like gene from Bordetella pertussis:
RT homology to the fim2 gene.";

```



```

RL Mol. Microbiol. 2:539-543 (1988).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=Tohama I / ATCC BAA-589 / NCTC 13251;
RX MEDLINE=22827954; PubMed=12910271;
RA Parkhill J., Sebahia M., Preston A., Murphy L.D., Thomson N.,
RA Harris D.E., Holden M.T.G., Churcher C.M., Bentley S.D., Mungall K.L.,
RA Cordero-Tarraga A.M., Temple L., James K., Harris B., Quail M.A.,
RA Achtman M., Atkin R., Baker S., Basham D., Bason N., Cherevach I.,
RA Chillingworth T., Collins M., Cronin A., Davis P., Doggett J.,
RA Feltwell T., Goble A., Hamlin N., Hauser H., Holtroyd S., Jagels K.,
RA Leather S., Moulé S., Norberczak H., O'Neill S., Ormond D., Price C.,
RA Rabinowitz E., Rutter S., Sanders M., Saunders D., Seeger K.,
RA Sharp S., Simmonds M., Skelton J., Squares R., Squares S., Stevens K.,
RA Unwin L., Whitehead S., Barrell B.G., Maskell D.J.;
RT "Comparative analysis of the genome sequences of Bordetella pertussis,
RL Nat. Genet. 35:32-40 (2003).
CC -!- SUBCELLULAR LOCATION: Pili structure on the cell surface.
CC -!- DISEASE: Bordetella pertussis is the causative agent of whooping
CC cough. An essential step in the disease process is the attachment
CC of the bacteria to the ciliated epithelium of the respiratory
CC tract, enabling the organism to resist normal host-clearance
CC mechanisms. It is unclear which bacterial cell surface component
CC are responsible for adherence but the fimbriae of B.pertussis are
CC prime candidates for being involved in this process.
CC -!- SIMILARITY: BELONGS TO THE FIMB/PAPA FAMILY OF FIMBRIA PROTEINS.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; Y00556; CA68634.1; -.
DR EMBL; AX5464; CA01770.1; -.
DR EMBL; B2640419; CAB42950.1; -.
DR PIR; S01929; S01929.
DR InterPro; IPR008966; Adhes bact.
DR Pfam; PF00419; Fimbrial; 1.
KW Fimbria; Signal; Complete proteome.
FT SIGNAL 1 21
FT CHAIN 22 201 FIMBRIAL PROTEIN FIMX.
FT DISULFID 37 79 PROBABLE.
FT CONFLICT 165 166 AK -> GQ (IN REF. 1).
FT SEQUENCE 201 AA; 21462 MW; F47670CC055B03F6 CRC64;

Query Match 1.5%; Score 8; DB 1; Length 201;
Best Local Similarity 100.0%; Pred. No. 4.3;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 10 LLGAALAG 17
Db 7 LLGAALAG 14

RESULT 11
Y726 METJA
AC Q58136;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Hypothetical protein W07026.
GN MJ0726.
OS Methanococcus jannaschii.
OC Archaea; Euryarchaeota; Methanococci; Methanococcales;
OC Methanocaldococcaceae; Methanocaldococcus.
OX NCBI_TaxID=2190;
RN [1]

```

```

RP SEQUENCE FROM N.A.
RC STRAIN=JAL-1 / DSM 2661 / ATCC 43067;
RX MEDLINE=96337999; PubMed=8688087;
RA Bult C.J., White O., Olsen G.J., Zhou L., Fleischmann R.D.,
RA Sutton G.G., Blake J.A., FitzGerald L.M., Clayton R.A., Gocayne J.D.,
RA Kleravage A.R., Dougherty B.A., Tomb J.-F., Adams M.D., Reich C.I.,
RA Overbeek R., Kirchges N.E.F., Weinstock K.G., Merrick J.M., Glöck A.,
RA Scott J.L., Geoghagen N.S.M., Weidman J.F., Fuhrmann J.L., Nguyen D.,
RA Utterback T.R., Kelley J.M., Peterson J.D., Sadow P.W., Hanna M.C.,
RA Cotton M.D., Roberts K.M., Hurst M.A., Kaine B.P., Borodovsky M.,
RA Klenk H.-P., Fraser C.M., Smith H.O., Woese C.R., Venter J.C.;
RT "Complete genome sequence of the methanogenic archaeon, Methanococcus
RN jannaschii."
RL Science 273:1058-1073 (1996).
CC -!- SIMILARITY: SOME, TO COENZYME F420 HYDROGENASE GAMMA SUBUNIT.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; U67519; AAB98722.1; -.
DR PIR; F64390; F64390.
DR TIGR; MJ0726; -.
DR InterPro; IPR001450; 4Fe4S ferredoxin.
DR InterPro; IPR006137; Oxidored_q6.
DR Pfam; PF00037; fer4; 1.
DR Pfam; PF01058; Oxidored_q6; 1.
DR PROSITE; PS00198; 4FE4S FERREDOXIN; 1.
KW Hypothetical protein; Oxidoreductase; Iron-sulfur; 4Fe-4S;
KW Electron transport; Complete proteome.
FT METAL 169 169 IRON-SULFUR 1 (4FE-4S) (POTENTIAL).
FT METAL 172 172 IRON-SULFUR 1 (4FE-4S) (POTENTIAL).
FT METAL 175 175 IRON-SULFUR 1 (4FE-4S) (POTENTIAL).
FT METAL 179 179 IRON-SULFUR 2 (4FE-4S) (POTENTIAL).
FT METAL 197 197 IRON-SULFUR 2 (4FE-4S) (POTENTIAL).
FT METAL 200 200 IRON-SULFUR 2 (4FE-4S) (POTENTIAL).
FT METAL 203 203 IRON-SULFUR 2 (4FE-4S) (POTENTIAL).
FT METAL 207 207 IRON-SULFUR 1 (4FE-4S) (POTENTIAL).
FT SEQUENCE 216 AA; 23789 MW; 2FD9F067698982FE CRC64;

Query Match 1.5%; Score 8; DB 1; Length 216;
Best Local Similarity 100.0%; Pred. No. 4.6;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 107 KEIVDSYL 114
Db 117 KEIVDSYL 124

RESULT 12
BOR4 MOUSE
ID BOR4 MOUSE STANDARD; PRT; 349 AA.
AC Q9JMS6; Q9QZT8;
DT 28-FEB-2003 (Rel. 41, Created).
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Cdc42 effector protein 4 (Binder of Rho GTPase 4).
GN CDC42BP4 OR BORG4 OR CEP4.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A., AND TISSUE SPECIFICITY.
RC TISSUE=ileal mucosa;
RX MEDLINE=21036164; PubMed=11185749;
RA Osada N., Kusuda J., Suzuki Y., Sugano S., Hashimoto K.;
RT "Sequence analysis, gene expression, and chromosomal assignment of
RN mouse Borg4 gene and its human orthologue."

```

```

J. Hum. Genet. 45:374-377(2000).
[2]
RN RP SEQUENCE FROM N.A.
RX MEDLINE=22388257; PubMed=12477932;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek A.M., Gay L.J., Hulyk S.W.,
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahy J., Helton E., Kettelman M., Madan A., Rodrigues S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butterfield Y.S.N., Krzywinski M.I., Skaleka U., Smailus D.E.,
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length
RT human and mouse cDNA sequences";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [3]
RN RP SEQUENCE OF 1-143 FROM N.A., AND INTERACTION WITH TC10 AND CDC42.
RX MEDLINE=99421943; PubMed=10490598;
RA Joberty G., Perlungher R.R., Macata I.G.;
RT "The Borgs, a new family of Cdc42 and TC10 GTPase-interacting
RT proteins";
RL Mol. Cell. Biol. 19:6585-6597(1999).
CC -!- FUNCTION: Probably involved in the organization of the actin
CC cytoskeleton. May act downstream of CDC42 to induce actin
CC filament assembly leading to cell shape changes. Induces
CC pseudopodia formation, when overexpressed in fibroblasts.
CC -!- SUBUNIT: Interacts with CDC42 and TC10 in a GTP-dependent manner.
CC -!- SUBCELLULAR LOCATION: Cytoplasmic; membrane-associated (By
CC similarity).
CC -!- TISSUE SPECIFICITY: Ubiquitous.
CC -!- SIMILARITY: BELONGS TO THE BORG/CEP FAMILY.
CC -!- SIMILARITY: Contains 1 CRIB domain.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AB035088; BAA95932.1; -
DR EMBL; BC003857; AAH03857.1; -
DR EMBL; AF165114; AA047822.1; -
DR MGD; MGI:1923760; Cdc42ep4.
DR GO; GO:0005737; Cytoplasm; IDA.
DR GO; GO:0005515; F-protein binding; IPI.
DR GO; GO:0007266; P-Rho protein signal transduction; IDA.
DR InterPro; IPR000095; PAKBox/RhoBndg.
DR Pfam; PF00786; PBD; 1
DR PROSITE; PS50108; CRIB; 1.
DR CYTOSKELETON.
DR DOMAIN 27 41 CRIB.
DR CONFLICT 143 143
DR SEQUENCE 349 AA; 37869 MW; 7C44125A7083E16B CRC64;

Query Match 1.5%; Score 8; DB 1; Length 349;
Best Local Similarity 100.0%; Pred. No. 7;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 347 SKLPSLS 354
Db 131 SKLPSLS 138
```

```

RESULT 13
BOR4 HUMAN STANDARD; PRT; 356 AA.
AC Q9H3Q0; Q95828; Q96FT3;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Cdc42 effector protein 4 (Binder of Rho GTPase 4).
GN CDC42EP4 OR BOR4 OR CEP4.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RN RP SEQUENCE FROM N.A.
RX MEDLINE=21036164; PubMed=11185749;
RA Osada N., Kusuda J., Suzuki Y., Sugano S., Hashimoto K.;
RT "Sequence analysis, gene expression, and chromosomal assignment of
RT mouse Borg4 gene and its human orthologue.";
RL J. Hum. Genet. 45:374-377(2000).
RN [2]
RN RP SEQUENCE FROM N.A., AND CHARACTERIZATION.
RX TISSUE=Eye;
RX MEDLINE=21125609; PubMed=11035016;
RA Hirsch D.S., Pirone D.M., Burbelo P.D.;
RT "A new family of cdc42 effector proteins, CEPs, function in fibroblast
RT and epithelial cell shape changes.";
RL J. Biol. Chem. 276:875-883(2001).
RN [3]
RN RP SEQUENCE FROM N.A.
RX TISSUE=Brain;
RX MEDLINE=22388257; PubMed=12477932;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek A.M., Gay L.J., Hulyk S.W.,
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahy J., Helton E., Kettelman M., Madan A., Rodrigues S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Butterfield Y.S.N., Krzywinski M.I., Skaleka U., Smailus D.E.,
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length
RT human and mouse cDNA sequences";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
CC -!- FUNCTION: Probably involved in the organization of the actin
CC cytoskeleton. May act downstream of CDC42 to induce actin
CC filament assembly leading to cell shape changes. Induces
CC pseudopodia formation, when overexpressed in fibroblasts.
CC -!- SUBUNIT: Interacts with CDC42 and TC10 in a GTP-dependent manner
CC (By similarity).
CC -!- SUBCELLULAR LOCATION: Cytoplasmic; membrane-associated.
CC -!- TISSUE SPECIFICITY: Not detected in any of the adult tissues
CC tested. May be expressed only in fetal or embryonic tissues.
CC -!- SIMILARITY: BELONGS TO THE BORG/CEP FAMILY.
CC -!- SIMILARITY: Contains 1 CRIB domain.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
```

```
CC -----
DR EMBL; AB042237; BAB1272.1; -
DR EMBL; AF099664; AAD16299.1; -
DR EMBL; BC010451; AAH10451.1; -
DR Genbank; HGNC:17147; CDC42EP4.
DR MM; 605468; -
DR InterPro; IPR000095; PAKbox/RhoBndng.
DR Pfam; PF00786; PBD; 1.
DR PROSITE; PS0108; CRIB; 1.
KW Cytoskeleton.
FT DOMAIN 27 41 CRIB.
FT DOMAIN 70 74 POLY-SER.
FT CONFLICT 2 2 P -> L (IN REF. 3).
FT CONFLICT 33 33 D -> T (IN REF. 2).
FT CONFLICT 288 288 A -> T (IN REF. 2).
FT CONFLICT 296 296 A -> T (IN REF. 2).
FT CONFLICT 339 339 P -> PP (IN REF. 2).
FT CONFLICT 355 355 MISSING (IN REF. 2).
SQ SEQUENCE 356 AA; 37979 MW; 2CF677C60C6EF1B5 CRC64;

Query Match 1.5%; Score 8; DB 1; Length 356;
Best Local Similarity 100.0%; Pred. No. 7.1;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 347 SKLPKSLSL 354
Db 133 SKLPKSLSL 140
|||||
133 SKLPKSLSL 140

RESULT 14
TCR3_ECOLI ID TCR3_ECOLI STANDARD; PRT; 396 AA.
AC P02981;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Tetracycline resistance protein, class C (TETA(C)).
GN TETA.
OS Escherichia coli.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Escherichia.
OX NCBI_TaxID=562;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=80002802; PubMed=383387;
RA Sutcliffe J.G.;
RT "Complete nucleotide sequence of the Escherichia coli plasmid pBR322.";
RL Cold Spring Harb. Symp. Quant. Biol. 43:77-90(1979).
RN [2]
RP REVISIONS, AND IDENTIFICATION OF PROTEIN.
RX MEDLINE=83117828; PubMed=6337373;
RA Livneh Z.;
RT "Directed mutagenesis method for analysis of mutagen specificity: application to ultraviolet-induced mutagenesis.";
RL Proc. Natl. Acad. Sci. U.S.A. 80:237-241(1983).
RN [3]
RP REVISIONS, AND IDENTIFICATION OF PROTEIN.
RX MEDLINE=83263146; PubMed=6307828;
RA Peden K.W.C.;
RT "Revised sequence of the tetracycline-resistance gene of pBR322.";
RL Gene 22:277-280(1983).
CC -1- FUNCTION: Resistance to tetracycline by an active tetracycline efflux. This is an energy-dependent process that decreases the accumulation of the antibiotic in whole cells. This protein functions as a metal-tetracycline/H(+) antiporter.
CC -1- SUBCELLULAR LOCATION: Integral membrane protein. Inner membrane.
CC -1- SIMILARITY: BELONGS TO THE MAJOR FACILITATOR FAMILY (ALSO KNOWN AS THE DRUG RESISTANCE TRANSLUCASE FAMILY).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation -
```

```
CC the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).
CC -----
DR EMBL; J01749; AAB59735.1; -
DR PIR; B90923; YTEC32.
DR InterPro; IPR007114; MFS.
DR InterPro; IPR005828; Sub_transporter.
DR InterPro; IPR005829; Sug_transporter.
DR InterPro; IPR001958; TCR_Teta.
DR Pfam; PF00083; sugar_tr; 1.
DR PRINTS; PRO1035; TCR_TETA.
DR PROSITE; PS00850; MFS; 1.
DR PROSITE; PS00216; SUGAR_TRANSPORT_1; UNKNOWN 1.
KW Antibiotic resistance; Transmembrane; Inner membrane; Transport;
KW Antiport; Ion transport; Hydrogen ion transport; Plasmid.
FT TRANSMEM 8 28 POTENTIAL.
FT TRANSMEM 44 64 POTENTIAL.
FT TRANSMEM 76 96 POTENTIAL.
FT TRANSMEM 104 124 POTENTIAL.
FT TRANSMEM 133 153 POTENTIAL.
FT TRANSMEM 161 181 POTENTIAL.
FT TRANSMEM 216 236 POTENTIAL.
FT TRANSMEM 246 267 POTENTIAL.
FT TRANSMEM 279 299 POTENTIAL.
FT TRANSMEM 300 320 POTENTIAL.
FT TRANSMEM 338 358 POTENTIAL.
FT TRANSMEM 365 385 POTENTIAL.
SQ SEQUENCE 396 AA; 41510 MW; 0B9C82B811A2F7DF CRC64;

Query Match 1.5%; Score 8; DB 1; Length 396;
Best Local Similarity 100.0%; Pred. No. 7.8;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 6 LLASLLGA 13
Db 76 LLASLLGA 83
|||||
76 LLASLLGA 83

RESULT 15
MNTC_BACSU ID MNTC_BACSU STANDARD; PRT; 435 AA.
AC O35024;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Manganese transport system membrane protein mntC.
GN MNTC OR BSU30750.
OS Bacillus subtilis.
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
OX NCBI_TaxID=1423;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=168;
RX MEDLINE=98048467; PubMed=9387221;
RA Lapidus A., Galleron N., Sorokin A., Ehrlich S.D.;
RT "Sequencing and functional annotation of the Bacillus subtilis genes in the 200 kb rmb-dnaB region.";
RL Microbiology 143:3431-3441(1997).
RN [2]
RP SEQUENCE FROM N.A.
RX STRAIN=168;
RX MEDLINE=98044033; PubMed=9384377;
RA Kunst F., Ogasawara N., Moszer I., Albertini A.M., Alloni G., Azevedo V., Bertero M.G., Bessieres P., Bolotin A., Borchert S., Borries R., Boursier L., Brans A., Braun M., Brignell S.C., Bron S., Brouillet S., Bruschi C.V., Caldwell B., Capuano V., Carter N.M., Choi S.K., Codani J.J., Connerton I.F., Cummings N.J., Daniel R.A., Denizot F., Devine K.W., Dusterhoft A., Ehrlich S.D., Emerson P.T., Entian K.D., Errington J., Fabret C., Ferrari E., Foulger D., Fritz C., Fujita M., Fujita Y., Fuma S., Galizzi A., Galleron N.,
```

Search completed: May 5, 2004, 13:46:14  
Job time : 11.5605 secs

```
RA Ghim S.Y., Glaser P., Goffeau A., Gollightly E.J., Grandi G.,
RA Giuseppe G., Guy B.J., Haga K., Haiech J., Harwood C.R., Henaut A.,
RA Hilbert H., Holsappel S., Hoono S., Hullo M.F., Itaya M., Jones L.,
RA Joris B., Karamata D., Kasahara Y., Kjaer-Blanchard M., Klein C.,
RA Kobayashi Y., Koetter P., Konigstein G., Krogh S., Kumano M.,
RA Kurita K., Lapidus A., Lardinois S., Lauber J., Lazarevic V.,
RA Lee S.M., Levine A., Liu H., Masuda S., Mauel C., Medigue C.,
RA Medina N., Mellado R.P., Mizuno M., Moestl D., Nakai S., Noback M.,
RA Noone D., O'Reilly M., Ogawa K., Ogiwara A., Oudega B., Park S.H.,
RA Parro V., Pohl T.M., Portetalle D., Porwollik S., Prescott A.M.,
RA Prescan E., Pujic P., Purnelle B., Rapoport G., Rey M., Reynolds S.,
RA Rieger M., Rivoita C., Roche E., Roche B., Rose M., Sadaie Y.,
RA Sato T., Scanlan E., Schleich S., Schroeter R., Scoffone F.,
RA Sekiguchi J., Sekowska A., Seror S.J., Serror P., Shin B.S., Soldo B.,
RA Sorokin A., Tacconi E., Takagi T., Takahashi H., Takemaru K.,
RA Takeuchi M., Tamakoshi A., Tanaka T., Terpstra P., Tognoni A.,
RA Viari A., Wambutt R., Wedler E., Wedler H., Weitzenecker T.,
RA Winters P., Wipat A., Yamamoto H., Yamane K., Yasumoto K., Yata K.,
RA Yoshida K., Yoshikawa H.F., Zumstein E., Yoshikawa H., Danchin A.;
RT "The complete genome sequence of the Gram-positive bacterium Bacillus
RT subtilis."
RL Nature 390:249-256(1997).
RN [3]
RP POSSIBLE FUNCTION.
RX MEDLINE=20223631; PubMed=10760146;
RA Que O., Helmann J.D.;
RT "Manganese homeostasis in Bacillus subtilis is regulated by MntR, a
RT bifunctional regulator related to the diphtheria toxin repressor
RL Mol Microbiol. 35:1454-1468(2000).
CC
CC -!- FUNCTION: This protein is probably a component of a manganese
CC permease, a binding protein-dependent, ATP-driven transport
CC system.
CC -!- SUBCELLULAR LOCATION: Integral membrane protein (Probable).
CC -!- SIMILARITY: BELONGS TO THE ABC-3 SUBFAMILY OF INTEGRAL MEMBRANE
CC PROTEINS.
CC
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AF008220; AAC00231.1; -
DR EMBL; Z99119; CAB15053.1; -
DR PIR; D69992; D69992.
DR Subtilist; BG13053; mntC.
DR InterPro; IPR001626; ABC_transpt3.
DR Pfam; PF00950; ABC-3; 1.
KW Transport; Ion transport; Transmembrane; Complete proteome.
FT TRANSMEM 17 37
FT TRANSMEM 42 62 POTENTIAL.
FT TRANSMEM 68 88 POTENTIAL.
FT TRANSMEM 98 118 POTENTIAL.
FT TRANSMEM 143 163 POTENTIAL.
FT TRANSMEM 166 186 POTENTIAL.
FT TRANSMEM 189 209 POTENTIAL.
FT TRANSMEM 228 248 POTENTIAL.
FT TRANSMEM 255 275 POTENTIAL.
SQ SEQUENCE 435 AA; 47944 MW; 68E6D590D90A73BC CRC64;

Query Match 1.5%; Score 8; DB 1; Length 435;
Best Local Similarity 100.0%; Pred. No. 8.5;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 10 LLGAALAG 17
DB 71 LLGAALAG 78
```

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 5, 2004, 13:42:09 ; Search time 16.595 Seconds  
(without alignments)  
3037.320 Million cell updates/sec

Title: US-09-743-684A-1

Perfect score: 524

Sequence: 1 MYALFLASLLGAALAGPVL.....NTETAACNAVECHKRHVWN 524

Scoring table: OLIGO

Gapop 60.0 , Gapext 60.0

Searched: 283366 seqs, 96191526 residues

Word size : 0

Total number of hits satisfying chosen parameters: 283366

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database :

PIR 78.\*

1: pir1.\*

2: pir2.\*

3: pir3.\*

4: pir4.\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	284	54.2	527	1 SAHUP	saposin precursor
2	22	4.2	79	2 A49475	cerebroside sulfat
3	22	4.2	554	1 A28716	saposin precursor
4	22	4.2	557	1 JH0604	saposin precursor
5	13	2.5	80	2 S21770	saposin-C - bovine
6	9	1.7	81	2 A32026	glucosylceramide b
7	8	1.5	162	2 AD1758	hypothetical prote
8	8	1.5	177	2 T47133	hypothetical prote
9	8	1.5	186	1 HHFP23	heat shock protein
10	8	1.5	186	2 B20647	heat shock protein
11	8	1.5	201	2 S36451	fimbrial protein f
12	8	1.5	201	2 S01929	fimbrial protein p
13	8	1.5	215	2 S56540	fim1 protein - Esc
14	8	1.5	215	2 E86129	fimbrial protein (
15	8	1.5	215	2 B91288	fimbrial protein (
16	8	1.5	216	2 F64390	coenzyme F420 hydr
17	8	1.5	216	2 T48482	hypothetical prote
18	8	1.5	225	2 T02241	probable germin pr
19	8	1.5	274	4 JC5585	hypothetical mutan
20	8	1.5	326	2 B90841	probable ABC trans
21	8	1.5	326	2 A85699	probable ABC trans
22	8	1.5	354	2 AP3614	probable ABC trans
23	8	1.5	396	1 YTEC32	tetracycline resis
24	8	1.5	435	2 D69992	ABC transporter (m
25	8	1.5	655	2 D85988	hypothetical prote
26	8	1.5	655	2 B65116	hypothetical prote
27	8	1.5	655	2 A91143	hypothetical prote
28	8	1.5	813	1 A49123	fibroblast growth
29	7	1.3	62	2 S77874	probable preprotei

#### RESULT 1

##### SAHUP

saposin precursor [validated] - human

N;Alternate names: cerebroside sulfate activator protein; co-beta-glucosidase; component  
ein (SAP); sphingolipid activator protein A2; sulfate sulfatase activator protein  
N;Contains: prosaposin; saposin A; saposin B; saposin C; saposin D  
C;Species: Homo sapiens (man)

C;Date: 30-Jun-1992 #sequence\_revision 17-Nov-1995 #text\_change 08-Dec-2000  
C;Accession: JX0061; A57368; A42003; B42003; C42003; D42003; A30367; S34740; S36140; S36  
0226; I37265; I37264

J;Nakano, T.; Sandhoff, K.; Stuenkel, J.; Christomanou, H.; Suzuki, K.

A;Title: Structure of full-length cDNA coding for sulfate activator, a Co-beta-glucosi-

A;Reference number: JX0061; MUID:89255151; PMID:2498298

A;Accession: JX0061

A;Molecule type: mRNA

A;Residues: 1-527 <NAK>

A;Cross-references: GB:D00422; NID:g220063; PIDN:BAA00321.1; PID:g220064

A;Note: alternative splice form 1

A;Accession: A57368

A;Molecule type: mRNA

A;Residues: 1-259,263-527 <NA2>

A;Cross-references: GB:J03015; GB:J03086; NID:g337755; PIDN:AAB59494.1; PID:g337756

A;Note: alternative splice form 2

R;Rorman, E.G.; Scheinker, V.; Grabowski, G.A.

Genomics 13, 312-318, 1992

A;Title: Structure and evolution of the human prosaposin chromosomal gene.

A;Reference number: A42003; MUID:92307663; PMID:1612590

A;Accession: A42003

A;Molecule type: DNA

A;Residues: 50-140 <ROR>

A;Cross-references: GB:M86181

A;Note: sequence extracted from NCBI backbone (NCBIN:107235, NCBIP:107236)

A;Accession: B42003

A;Molecule type: DNA

A;Residues: 185-259,263-276 <RO2>

A;Note: sequence extracted from NCBI backbone (NCBIN:107235, NCBIP:107237)

A;Accession: C42003

A;Molecule type: DNA

A;Residues: 305-393 <RO3>

A;Note: sequence extracted from NCBI backbone (NCBIN:107235, NCBIP:107238); sequence inc

A;Accession: D42003

A;Molecule type: DNA

A;Residues: 399-487 <RO4>

A;Note: sequence extracted from NCBI backbone (NCBIN:107235, NCBIP:107239); sequence inc

R;Rorman, E.G.; Grabowski, G.A.

Genomics 5, 486-492, 1989

A;Title: Molecular cloning of a human co-beta-glucosidase cDNA: evidence that four sphin

A;Reference number: A30367; MUID:90129043; PMID:2515150

A;Accession: A30367

A;Molecule type: mRNA

A;Residues: 1-259,263-527 <RO5>

A;Cross-references: GB:J03077; NID:gl83230; PIDN:AAA52560.1; PID:gl83231  
A;Note: alternative splice form 2  
R;Hiraiwa, M.; O'Brien, J.S.; Kishimoto, Y.; Galdick, M.; Fluharty, A.L.; Ginns, E.I.; Arch. Biochem. Biophys. 304, 110-116, 1993  
A;Title: Isolation, characterization, and proteolysis of human prosaposin, the precursor of saposin A  
A;Reference number: S34740; MUID:93311991; PMID:8323276  
A;Accession: S34740  
A;Molecule type: protein  
A;Residues: 17-24;165-172;180-189;301-305 <HIR>  
R;Tyynelae, J.; Palmer, D.N.; Baumann, M.; Haltia, M. FEBS Lett. 330, 8-12, 1993  
A;Title: Storage of saposins A and D in infantile neuronal ceroid-lipofuscinosis.  
A;Reference number: S36140; MUID:93380576; PMID:8370464  
A;Accession: S36140  
A;Molecule type: protein  
A;Residues: 'XX', 62, 'X', 64-65, 'X', 67-79, 'X', 81-84 <TY>  
A;Note: saposin A  
A;Accession: S36141  
A;Molecule type: protein  
A;Residues: 'XX', 413-414, 'X', 416-428, 'X', 430-434 <TY2>  
A;Note: saposin D  
R;Holtzman, H.; Sandhoff, K.; Kwon, H.Y.; Harzer, K.; Nakano, T.; Suzuki, K. J. Biol. Chem. 266, 7556-7560, 1991  
A;Title: Sulfatide activator protein. Alternative splicing that generates three mRNAs and a protein  
A;Reference number: S36988; MUID:91210267; PMID:2019586  
A;Accession: S36988  
A;Status: nucleic acid sequence not shown; translation not shown  
A;Molecule type: mRNA  
A;Residues: 1-240, 'S', 242-527 <HOL>  
A;Cross-references: EMBL:M60255; NID:g337759; PIDN:AAA36594.1; PID:g337760  
A;Note: the nucleotide sequence was submitted to the EMBL Data Library, June 1991  
A;Note: cerebroside sulfate activator protein mutant MU-9; corresponds to alternative splicing  
A;Accession: S36989  
A;Status: nucleic acid sequence not shown; translation not shown  
A;Molecule type: mRNA  
A;Residues: 1-240, 'S', 242-527, 263-527 <HO2>  
A;Cross-references: EMBL:M60257; NID:g337764; PIDN:AAA36595.1; PID:g337765  
A;Note: the nucleotide sequence was submitted to the EMBL Data Library, June 1991  
A;Note: cerebroside sulfate activator protein mutant MU-0; corresponds to alternative splicing  
A;Accession: S36990  
A;Status: nucleic acid sequence not shown; translation not shown  
A;Molecule type: mRNA  
A;Residues: 1-240, 'S', 242-259, 261-527 <HO3>  
A;Cross-references: EMBL:M60258; NID:g337766; PIDN:AAA36596.1; PID:g337767  
A;Note: the nucleotide sequence was submitted to the EMBL Data Library, June 1991  
A;Note: cerebroside sulfate activator protein mutant MU-6; corresponds to alternative splicing  
R;Kondoh, K.; Hinenio, T.; Sano, A.; Kakimoto, Y. Biochem. Biophys. Res. Commun. 181, 286-292, 1991  
A;Title: Isolation and characterization of prosaposin from human milk.  
A;Reference number: PS0330; MUID:92068206; PMID:1958198  
A;Accession: PS0330  
A;Molecule type: protein  
A;Residues: 17-24, 'X', 26 <KON>  
A;Experimental source: milk  
R;Kretz, K.A.; Carson, G.S.; Morimoto, S.; Kishimoto, Y.; Fluharty, A.L.; O'Brien, J.S. Proc. Natl. Acad. Sci. U.S.A. 87, 2541-2544, 1990  
A;Title: Characterization of a mutation in a family with saposin B deficiency: a glycosylation defect  
A;Reference number: A35985; MUID:90207231; PMID:2320574  
A;Accession: A35985  
A;Molecule type: mRNA  
A;Residues: 213-221 <KRE>  
A;Cross-references: GB:M32221  
A;Accession: B35985  
A;Status: nucleic acid sequence not shown; translation not shown  
A;Molecule type: mRNA  
A;Residues: 1-259, 263-527 <KR2>  
A;Cross-references: GB:M32221; NID:g337761; PIDN:AAA60303.1; PID:g337762  
A;Experimental source: lymphoblast  
A;Accession: C35985  
A;Molecule type: mRNA  
A;Residues: 213-216, 'I', 218-221 <KR3>  
A;Note: sequence from patients with activator-deficient metachromatic leukodystrophy; th  
R;Fuerst, W.; Schubert, J.; Machleidt, W.; Meyer, H.B.; Sandhoff, K.

Eur. J. Biochem. 192, 709-714, 1990  
A;Title: The complete amino-acid sequences of human ganglioside GM2 activator protein and its precursor  
A;Reference number: S13195; MUID:91006185; PMID:2209618  
A;Accession: S13196  
A;Molecule type: protein  
A;Residues: 195-259, 263-277 <FUE>  
R;Morimoto, S.; Martin, B.M.; Yamamoto, Y.; Kretz, K.A.; O'Brien, J.S.; Kishimoto, Y. Proc. Natl. Acad. Sci. U.S.A. 86, 3389-3393, 1989  
A;Title: Saposin A: second cerebrosidase activator protein.  
A;Reference number: A32784; MUID:89240739; PMID:2717620  
A;Accession: A32784  
A;Molecule type: protein  
A;Residues: 60-84;86-107;109-119;125-134 <MOR>  
R;O'Brien, J.S.; Kretz, K.A.; Dewji, N.; Wenger, D.A.; Esch, F.; Fluharty, A.L. Science 241, 1098-1101, 1988  
A;Title: Coding of two sphingolipid activator proteins (SAP-1 and SAP-2) by same genetic locus  
A;Reference number: A41240; MUID:88321660; PMID:2842863  
A;Accession: A41240  
A;Molecule type: mRNA  
A;Residues: 'GSSR', 18-259, 263-299, 'D', 301-302, 'D', 304-527 <OAB>  
A;Cross-references: GB:J03086  
R;Dewji, N.N.; Wenger, D.A.; O'Brien, J.S. Proc. Natl. Acad. Sci. U.S.A. 84, 8652-8656, 1987  
A;Title: Nucleotide sequence of cloned cDNA for human sphingolipid activator protein 1 p  
A;Reference number: S02289; MUID:88068647; PMID:2825202  
A;Accession: S02289  
A;Status: significant sequence differences  
A;Molecule type: mRNA  
A;Cross-references: EMBL:J03015  
A;Note: this sequence corrected by A41240  
R;Kleinschmidt, T.; Christomanou, H.; Braunitzer, G. Biol. Chem. Hoppe-Seyler 369, 1361-1365, 1988  
A;Title: Complete amino-acid sequence of the naturally occurring A(2) activator protein 1 p  
A;Reference number: S02028; MUID:89207118; PMID:3242555  
A;Accession: S02028  
A;Molecule type: protein  
R;Fuerst, W.; Machleidt, W.; Sandhoff, K. Biol. Chem. Hoppe-Seyler 369, 317-328, 1988  
A;Title: The precursor of sulfatide activator protein is processed to three different products  
A;Reference number: S00813; MUID:89000190; PMID:3048308  
A;Accession: S00813  
A;Molecule type: protein  
R;Kleinschmidt, T.; Christomanou, H.; Braunitzer, G. Biol. Chem. Hoppe-Seyler 368, 1571-1578, 1987  
A;Title: Complete amino-acid sequence and carbohydrate content of the naturally occurring A(2) activator protein 1 p  
A;Reference number: S00226; MUID:88163077; PMID:3442600  
A;Accession: S00226  
A;Molecule type: protein  
A;Residues: 314-393 <KL2>  
R;Vaccaro, A.M.; Salvioli, R.; Barca, A.; Tatti, M.; Ciaffoni, F.; Maras, B.; Siciliano, J. Biol. Chem. 270, 9953-9960, 1995  
A;Title: Structural analysis of saposin C and B. Complete localization of disulfide bridge  
A;Reference number: A57297; MUID:95247790; PMID:7730378  
A;Contents: annotation; disulfide bonds; glycosylation  
R;Holtzman, H.; Sandhoff, K.; Fuerst, W.; Kwon, H.Y.; Schnabel, D.; Suzuki, K. FEBS Lett. 280, 267-270, 1991  
A;Title: The organization of the gene for the human cerebroside sulfate activator protein  
A;Reference number: I37264; MUID:91192146; PMID:2013321  
A;Accession: I37265  
A;Status: translated from GB/EMBL/DBJ  
A;Molecule type: DNA  
A;Residues: 59-125 <RES>  
A;Cross-references: EMBL:X57107; NID:g30234; PIDN:CAA40391.1; PID:g30235  
A;Accession: I37264  
A;Status: translated from GB/EMBL/DBJ  
A;Molecule type: DNA  
A;Residues: 307-516 <RE2>  
A;Cross-references: EMBL:X57108; NID:g30232; PIDN:CAA40392.1; PID:gl565257  
A;Note: sequence revised relative to PID:g30233 (corrected coding region)  
C;Genetics:

A:Gene: GDB:PSAP; GLBA  
A:Cross-references: GDB:120366; OMIM:176801  
A:Map position: 10q22.1-10q22.1  
A:Introns: 83/3; 338/3; 401/1; 453/3; 480/3  
A:Note: defects in this gene may cause variant Gaucher disease, variant Tay-Sachs disease  
A:Note: list of introns is incomplete

Query Match 54.2%; Score 284; DB 1; Length 527;  
Best Local Similarity 99.4%; Pred. No. 1.7e-284;  
Matches 524; Conservative 0; Mismatches 0; Indels 3; Gaps 1;

Qy 1 MYALFLASLIGALAGVILGKCTRGSAVWCQNVKTASDCGAVKHCLOTVWNKPTVKS 60  
Db 1 MYALFLASLIGALAGVILGKCTRGSAVWCQNVKTASDCGAVKHCLOTVWNKPTVKS 60

Qy 61 LPCDICDVVTAAGDMLKDNATEEILVLEKTCOMLPKPNNSASCKEIVDSYLPVILI 120  
Db 61 LPCDICDVVTAAGDMLKDNATEEILVLEKTCOMLPKPNNSASCKEIVDSYLPVILI 120

Qy 121 IKGEMSRPGEVCSALNLCESLQKHLAEINHQLESNNKIPELDMTEVVAPFMANIPLLY 180  
Db 121 IKGEMSRPGEVCSALNLCESLQKHLAEINHQLESNNKIPELDMTEVVAPFMANIPLLY 180

Qy 181 PODGPRKPKQKNGDVQCDCIQWTDIQTAVRTNSTFVQALVHVKEECDRLGPGWADI 240  
Db 181 PODGPRKPKQKNGDVQCDCIQWTDIQTAVRTNSTFVQALVHVKEECDRLGPGWADI 240

Qy 241 CKNYISQYSEITAIOMMHM---QPEICALGVFCDEVKEMPQTLVPAKVASKNVI PALE 297  
Db 241 CKNYISQYSEITAIOMMHMQQQPEICALGVFCDEVKEMPQTLVPAKVASKNVI PALE 300

Qy 298 LVEPIKHEVPAKSDVYCEVCEFLVKEVTKLIDNNKTEKEITLDAFDKWCXKLPKSLSEEC 357  
Db 301 LVEPIKHEVPAKSDVYCEVCEFLVKEVTKLIDNNKTEKEITLDAFDKWCXKLPKSLSEEC 360

Qy 358 QEVVDTTGSSTLSILBEVSPVLSVCSHLCSGTRLPALTTHVHTQPKDGGCFCEVCKKLVG 417  
Db 361 QEVVDTTGSSTLSILBEVSPVLSVCSHLCSGTRLPALTTHVHTQPKDGGCFCEVCKKLVG 420

Qy 418 YLDRNLEKNSKQELTAALEKGCFLPDPYQKQCDQFVAEYEPVLIETLVEWMDPSFVCL 477  
Db 421 YLDRNLEKNSKQELTAALEKGCFLPDPYQKQCDQFVAEYEPVLIETLVEWMDPSFVCL 480

Qy 478 KIGACPSAHKPLLTGTEKICWGPSYWCQNTETAACNAVEHCXKRVWN 524  
Db 481 KIGACPSAHKPLLTGTEKICWGPSYWCQNTETAACNAVEHCXKRVWN 527

RESULT 2  
A49475  
cerebroside sulfate activator - pig  
C:Species: Sus scrofa domestica (domestic pig)  
C:Date: 24-Feb-1994 #sequence\_revision 18-Nov-1994 #text\_change 17-May-1996  
C:Accession: A49475  
R:Stevens, R.L.; Faull, K.F.; Conklin, K.A.; Green, B.N.; Fluharty, A.L.  
Biochemistry 32, 4051-4059, 1993  
A:Title: Porcine cerebroside sulfate activator: further structural characterization and  
A:Reference number: A49475; MUID:932229506; PMID:8471613  
A:Accession: A49475  
A:Status: preliminary  
A:Molecule type: protein  
A:Residues: 1-79 <STE>  
A:Experimental source: kidney  
A:Note: sequence extracted from NCBI backbone (NCBIP:129597)  
C:Superfamily: saposin; saposin repeat homology  
F:1-79/Domain: saposin repeat homology <SAP>

Query Match 4.2%; Score 22; DB 2; Length 79;  
Best Local Similarity 100.0%; Pred. No. 9.1e-15;  
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 241 CKNYISQYSEITAIOMMHMQPK 262  
|||||

Db 47 CKNYISQYSEITAIOMMHMQPK 68

RESULT 3  
A28716  
saposin precursor - rat  
N:Alternate names: cerebroside sulfate activator protein; co-beta-glucosidase; component  
ein (SAP); sphingolipid activator protein A2; sulfated glycoprotein 1; sulfatide sulfata  
N:Contains: prosaposin; saposin A; saposin B; saposin C; saposin D  
C:Species: Rattus norvegicus (Norway rat)  
C:Date: 10-Sep-1999 #sequence\_revision 10-Sep-1999 #text\_change 10-Sep-1999  
C:Accession: A28716  
R:Collard, M.W.; Sylvester, S.R.; Tsuruta, J.K.; Griswold, M.D.  
Biochemistry 27, 4557-4564, 1988  
A:Title: Biosynthesis and molecular cloning of sulfated glycoprotein 1 secreted by rat  
A:Reference number: A28716; MUID:89000647; PMID:3048385  
A:Accession: A28716  
A:Molecule type: mRNA  
A:Residues: 1-554 <COL>  
A:Cross-references: GB:M19936; NID:g206904; PIDN:AAA42136.1; PID:g206905  
A:Note: parts of this sequence, including the amino end of the mature protein, were dete  
C:Function:  
A:Description: saposins bind sphingolipids, form hydrophilic complexes and make them acc  
A:Pathway: sphingolipid catabolism  
A:Note: saposins A and C (SAP-2) activate hydrolysis of glucocerebroside by beta-glucosy  
A:Note: saposin B (SAP-1) activates hydrolysis of galactocerebroside sulfate by arylsulfa  
A:Note: saposin D activates hydrolysis of sphingomyelin by sphingomyelin phosphodiesterase  
C:Superfamily: saposin; saposin repeat homology  
C:Keywords: alternative splicing; glycoprotein; lysosomal storage disease; lysosome; sph  
F:1-16/Domain: signal sequence #status predicted <SIG>  
F:17-554/Product: prosaposin #status predicted <PRO>  
F:55-148/Domain: saposin repeat homology <SAP1>  
F:60-143/Product: saposin A #status predicted <SAPA>  
F:189-280/Domain: saposin repeat homology <SAP2>  
F:194-273/Product: saposin B #status predicted <SAB1>  
F:306-397/Domain: saposin repeat homology <SAP3>  
F:310-389/Product: saposin C #status predicted <SAPC>  
F:431-522/Domain: saposin repeat homology <SAP4>  
F:437-514/Product: saposin D #status predicted <SAPD>  
F:63-138,66-132,94-106,439-512,442-506,470-481/Disulfide bonds: #status predicted  
F:80,214,331,456/Binding site: carbonyl site: Asn (covalent) #status predicted  
F:197-270,200-264,229-240,314-387,317-381,345-356/Disulfide bonds: #status predicted

Query Match 4.2%; Score 22; DB 1; Length 554;  
Best Local Similarity 100.0%; Pred. No. 5.5e-14;  
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 431 EILAALEKGCFLPDPYQKQCD 452  
|||||  
Db 461 EILAALEKGCFLPDPYQKQCD 482

RESULT 4  
JH0604  
saposin precursor - mouse  
N:Alternate names: cerebroside sulfate activator protein; co-beta-glucosidase; component  
ein (SAP); sphingolipid activator protein A2; sulfated glycoprotein 1; sulfatide sulfata  
N:Contains: prosaposin; saposin A; saposin B; saposin C; saposin D  
C:Species: Mus musculus (house mouse)  
C:Date: 10-Sep-1999 #sequence\_revision 10-Sep-1999 #text\_change 10-Sep-1999  
C:Accession: JH0604  
R:Tsuda, M.; Sakiyama, T.; Endo, H.; Kitagawa, T.  
Biochem. Biophys. Res. Commun. 184, 1266-1272, 1992  
A:Title: The primary structure of mouse saposin.  
A:Reference number: JH0604; MUID:92272718; PMID:1590788  
A:Accession: JH0604  
A:Molecule type: mRNA  
A:Residues: 1-557 <TSU>  
A:Cross-references: GB:S36200; NID:g249386; PIDN:AAB22175.1; PID:g249387  
A:Experimental source: liver  
C:Function:  
A:Description: saposins bind sphingolipids, form hydrophilic complexes and make them acc  
A:Pathway: sphingolipid catabolism



A;Note: saposins A and C (SAP-2) activate hydrolysis of glucocerebroside by beta-glucosyl  
A;Note: saposin B (SAP-1) activates hydrolysis of galactocerebroside sulfate by arylsulfa  
C;Superfamily: saposin; saposin hydrolysis of sphingomyelin by sphingomyelin phosphodiesterase  
C;Keywords: alternative splicing; glycoprotein; lysosomal storage disease; lysosome; sph  
F;1-16/Domain: signal sequence #status predicted <SIG>  
F;17-557/Product: prosaposin #status predicted <PRO>  
F;55-148/Domain: saposin A #status predicted <SAP1>  
F;60-143/Product: saposin repeat homology <SAP2>  
F;194-276/Product: saposin B #status predicted <SAB1>  
F;309-400/Domain: saposin repeat homology <SAP3>  
F;313-392/Product: saposin C #status predicted <SAPC>  
F;434-525/Domain: saposin repeat homology <SAP4>  
F;440-517/Product: saposin D #status predicted <SAPD>  
F;63-138,66-132,94-106,197-273,200-267,229-240,317-390,320-384,348-359,442-515,445-509,4  
F;80,214,334,379,459/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 4.2%; Score 22; DB 1; Length 557;  
Best Local Similarity 100.0%; Pred. No. 5.6e-14;  
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 431 EILALEKGCSEFLPDYKQCD 452  
|||||  
Db 464 EILALEKGCSEFLPDYKQCD 485

RESULT 5  
S21770  
saposin-C - bovine  
C;Species: Bos primigenius taurus (cattle)  
C;Date: 19-Mar-1997 #sequence\_revision 19-Mar-1997 #text\_change 31-Oct-1997  
C;Accession: S21770  
R;Sano, A.; Mizuno, T.; Kondoh, K.; Hineno, T.; Ueno, S.; Kakimoto, Y.; Morita, N.  
Biochim. Biophys. Acta 1120, 75-80, 1992  
A;Title: Saposin-C from bovine spleen: complete amino acid sequence and relation between  
A;Reference number: S21770; MUID:92207994; PMID:1554743  
A;Accession: S21770  
A;Status: preliminary  
A;Molecule type: protein  
A;Residues: 1-80 <SAN>  
C;Superfamily: saposin; saposin repeat homology

Query Match 2.5%; Score 13; DB 2; Length 80;  
Best Local Similarity 100.0%; Pred. No. 1.8e-05;  
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 377 SPELVCSMLHCS 389  
|||||  
Db 67 SPELVCSMLHCS 79

RESULT 6  
A32026  
Glucosylceramide beta-glucosidase activator protein SAP-2 - guinea pig  
C;Species: Cavia porcellus (guinea pig)  
C;Date: 07-Jun-1990 #sequence\_revision 07-Jun-1990 #text\_change 17-May-1996  
C;Accession: A32026  
R;Sano, A.; Radin, N.S.; Johnson, L.L.; Tarr, G.E.  
J. Biol. Chem. 263, 19597-19601, 1988  
A;Title: The activator protein for glucosylceramide beta-glucosidase from guinea pig liv  
A;Reference number: A32026; MUID:89066787; PMID:3198642  
A;Accession: A32026  
A;Molecule type: protein  
A;Residues: 1-81 <SAN>  
C;Superfamily: saposin; saposin repeat homology  
F;1-81/Domain: saposin repeat homology <SAP>

Query Match 1.7%; Score 9; DB 2; Length 81;  
Best Local Similarity 100.0%; Pred. No. 0.25;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 357 CQEVDTYG 365

Db 47 CQEVDTYG 55  
|||||

## RESULT 7

AD1758  
hypothetical protein lin2609 [imported] - Listeria innocua (strain Clp11262)

C;Species: Listeria innocua  
C;Date: 27-Nov-2001 #sequence\_revision 27-Nov-2001 #text\_change 27-Nov-2001

C;Accession: AD1758  
R;Glaser, P.; Frangeul, L.; Buchrieser, C.; Amend, A.; Baquero, F.; Berche, P.; Bloecker,

; Dominguez-Bernal, G.; Duchaud, E.; Durand, L.; Dussurget, O.; Entian, K.D.; Feihl, H.

D.; Jones, L.M.; Karst, U.  
Science 294, 849-852, 2001

A;Authors: Kreft, J.; Kuhn, M.; Kunst, F.; Kurapkat, G.; Madueno, E.; Maitournam, A.; Mat

ok, C.; Schluter, T.; Simoes, N.; Tierrez, A.; Vazquez-Boland, J.A.; Voss, H.; Wehland,

A;Title: Comparative genomics of Listeria species.

A;Reference number: AB1077; MUID:21537279; PMID:11679669

A;Accession: AD1758

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-162 <GLA>

A;Cross-references: GB:AL592022; PIDN:CAC97836.1; PID:gi16415146; GSPDB:GN00178

A;Experimental source: strain Clp11262

C;Genetics:

A;Gene: lin2609

## Query Match

1.5%; Score 8; DB 2; Length 162;

Best Local Similarity 100.0%; Pred. No. 5;

Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 LFLASLL 11

|||||

Db 4 LFLASLL 11

## RESULT 8

T47133

hypothetical protein DKFZp761P19121.1 - human (fragment)

C;Species: Homo sapiens (man)

C;Date: 20-Apr-2000 #sequence\_revision 20-Apr-2000 #text\_change 20-Apr-2000

C;Accession: T47133

R;Poustka, A.; Wellenreuther, R.; Mewes, H.W.; Weil, B.; Wiemann, S.

submitted to the Protein Sequence Database, March 2000

A;Reference number: Z24374

A;Accession: T47133

A;Status: preliminary

A;Molecule type: mRNA

A;Residues: 1-177 <AAA>

A;Cross-references: EMBL:ALJ62013

A;Experimental source: adult amygdala; clone DKFZp761P19121

C;Genetics:

A;Note: DKFZp761P19121.1

## Query Match

1.5%; Score 8; DB 2; Length 177;

Best Local Similarity 100.0%; Pred. No. 5.5;

Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 431 EILALEK 438

|||||

Db 146 EILALEK 153

## RESULT 9

HHFF23

heat shock protein 23 - fruit fly (Drosophila melanogaster)

C;Species: Drosophila melanogaster

C;Date: 15-Nov-1984 #sequence\_revision 15-Nov-1984 #text\_change 16-Feb-1997

C;Accession: A02919

R;Ingolia, T.D.; Craig, E.A.

Proc. Natl. Acad. Sci. U.S.A. 79, 2360-2364, 1982

A;Title: Four small Drosophila heat shock proteins are related to each other and to mam

A;Reference number: A93909; MUID:82248004; PMID:6285380



A;Accession: A02919  
A;Molecule type: DNA  
A;Residues: 1-186 <ING>  
A;Note: the authors translated the codon CCC for residue 25 as Leu, CGA for residue 32 as  
C;Comment: This small heat shock protein is related to alpha crystallin.  
C;Genetics:  
A;Gene: FlyBase:Hsp23  
A;Cross-references: FlyBase:FBgn0001224  
A;Map position: 3L (67B)  
C;Superfamily: alpha-crystallin  
C;Keywords: heat shock; stress-induced protein

Query Match 1.5%; Score 8; DB 1; Length 186;  
Best Local Similarity 100.0%; Pred. No. 5.7;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 172 MANIPLLL 179  
Db 1 MANIPLLL 8

RESULT 10  
B20647  
heat shock protein 23 - fruit fly (*Drosophila melanogaster*)  
C;Species: *Drosophila melanogaster*  
C;Date: 09-Sep-1987 #sequence\_revision 09-Sep-1987 #text\_change 21-Jul-2000  
C;Accession: B20647  
R;Southgate, R.; Ayme, A.; Voellmy, R.  
J. Mol. Biol. 165, 35-57, 1983  
A;Title: Nucleotide sequence analysis of the *Drosophila* small heat shock gene cluster at  
A;Reference number: A92897; MUID:83189140; PMID:6302284  
A;Accession: B20647  
A;Molecule type: DNA  
A;Residues: 1-186 <SOU>  
A;Cross-references: GB:V00210; NID:98099; PIDN:CAA23494.1; PID:98100  
C;Genetics:  
A;Gene: FlyBase:Hsp23  
A;Cross-references: FlyBase:FBgn0001224  
C;Superfamily: alpha-crystallin

Query Match 1.5%; Score 8; DB 2; Length 186;  
Best Local Similarity 100.0%; Pred. No. 5.7;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 172 MANIPLLL 179  
Db 1 MANIPLLL 8

RESULT 11  
S36451  
fimbrial protein fimX - *Bordetella bronchiseptica*  
C;Species: *Bordetella bronchiseptica*  
C;Date: 06-Jan-1995 #sequence\_revision 06-Jan-1995 #text\_change 26-Aug-1999  
C;Accession: S36451  
R;Savelkoul, P.; de Kerf, D.; de Groot, L.; Willems, R.; Mooi, F.; van der Zeijst, B.; G  
submitted to the EMBL Data Library, February 1993  
A;Description: Characterization of fimX, a gene encoding a fimbrial subunit protein of B  
A;Reference number: S36451  
A;Accession: S36451  
A;Status: preliminary  
A;Molecule type: DNA  
A;Residues: 1-201 <SAV>  
A;Cross-references: EMBL:X74118; NID:g396788; PIDN:CAA52215.1; PID:g396789  
C;Genetics:  
A;Gene: fimX  
C;Superfamily: type 1 fimbrial protein

Query Match 1.5%; Score 8; DB 2; Length 201;  
Best Local Similarity 100.0%; Pred. No. 6.3;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 10 LLGAALAG 17  
Db 7 LLGAALAG 14

RESULT 12  
S01929  
fimbrial protein precursor - *Bordetella pertussis*  
C;Species: *Bordetella pertussis*  
C;Date: 18-Oct-1989 #sequence\_revision 18-Oct-1989 #text\_change 26-Aug-1999  
C;Accession: S01929  
R;Pedroni, P.; Riboli, B.; de Ferra, F.; Grandi, G.; Toma, S.; Arico, B.; Rappuoli, R.  
Mol. Microbiol. 2, 539-543, 1988  
A;Title: Cloning of a novel pilin-like gene from *Bordetella pertussis*: homology to the f  
A;Reference number: S01929; MUID:89013896; PMID:2902506  
A;Accession: S01929  
A;Molecule type: DNA  
A;Residues: 1-201 <FED>  
A;Cross-references: EMBL:Y00556; NID:g39750; PIDN:CAA68634.1; PID:g39751  
C;Genetics:  
A;Gene: fimX  
C;Superfamily: type 1 fimbrial protein  
F;1-21/Domain: signal sequence #status predicted <SIG>  
F;22-201/Product: fimbrial protein #status predicted <MAT>

Query Match 1.5%; Score 8; DB 2; Length 201;  
Best Local Similarity 100.0%; Pred. No. 6.2;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 10 LLGAALAG 17  
Db 7 LLGAALAG 14

RESULT 13  
S56540  
fimi protein - *Escherichia coli* (strain K-12)  
C;Species: *Escherichia coli*  
C;Date: 28-Oct-1995 #sequence\_revision 03-Nov-1995 #text\_change 01-Mar-2002  
C;Accession: S56540; E65245; S53066  
R;Burland, V.; Plunkett III, G.; Sofia, H.J.; Daniels, D.L.; Blattner, F.R.  
Nucleic Acids Res. 23, 2105-2119, 1995  
A;Title: Analysis of the *Escherichia coli* genome VI: DNA sequence of the region from 92.1  
A;Reference number: S56314; MUID:95334362; PMID:7610040  
A;Accession: S56540  
A;Status: preliminary; nucleic acid sequence not shown; translation not shown  
A;Molecule type: DNA  
A;Residues: 1-215 <BUR>  
A;Cross-references: EMBL:U14003; NID:g1263172; PIDN:AAA97211.1; PID:g537156  
A;Note: the nucleotide sequence was submitted to the EMBL Data Library, August 1994  
R;Blattner, F.R.; Plunkett III, G.; Bloch, C.A.; Perna, N.T.; Burland, V.; Riley, M.; Co  
A.; Rose, D.J.; Mau, B.; Shao, Y.  
Science 277, 1453-1462, 1997  
A;Title: The complete genome sequence of *Escherichia coli* K-12.  
A;Reference number: A64720; MUID:97426617; PMID:9278503  
A;Accession: E65245  
A;Status: preliminary; nucleic acid sequence not shown; translation not shown  
A;Molecule type: DNA  
A;Residues: 1-215 <BLAT>  
A;Cross-references: GB:AE000502; GB:U00096; NID:g23637374; PIDN:AAC77271.1; PID:g1790770;  
A;Experimental source: strain K-12, substrain MG1655  
R;Marc, D.; Dho-Moulin, M.  
submitted to the EMBL Data Library, September 1994  
A;Description: Hypervariable sites within the major subunit of type 1 fimbriae, identifi  
A;Reference number: S53063  
A;Accession: S53066  
A;Status: preliminary  
A;Molecule type: DNA  
A;Residues: 37-77, 'A', '79-107', 'S', '109-215 <MAR>  
A;Cross-references: EMBL:Z37500; NID:g732680; PIDN:CAA85728.1; PID:g732684  
C;Genetics:  
A;Gene: fimI  
A;Start codon: GTG

C;Superfamily: type 1 fimbrial protein

Query Match 1.5%; Score 8; DB 2; Length 215;  
Best Local Similarity 100.0%; Pred. No. 6.5;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 LFLLASLL 11  
|||||||  
Db 42 LFLLASLL 49

#### RESULT 14

E86129  
fimbrial protein [imported] - Escherichia coli (strain O157:H7, substrain EDL933)  
C;Species: Escherichia coli  
C;Date: 16-Feb-2001 #sequence\_revision 16-Feb-2001 #text\_change 17-May-2002  
C;Accession: E86129  
R;Perna, N.T.; Plunkett III, G.; Burland, V.; Mau, B.; Glasner, J.D.; Rose, D.J.; Mayhew  
iller, L.; Grotbeck, E.J.; Davis, N.W.; Lim, A.; Dimalanta, E.; Potamouisis, K.; Apodaca,  
Nature 409, 529-533, 2001  
A;Title: Genome sequence of enterohemorrhagic Escherichia coli O157:H7.  
A;Reference number: A85480; MUID:21074935; PMID:11206551  
A;Accession: E86129  
A;Status: preliminary  
A;Molecule type: DNA  
A;Residues: 1-215 <STO>  
A;Cross-references: GB:AE005174; NID:G12519322; PIDN:AAG59497.1; GSPDB:GN00145; UWGP:Z59  
A;Experimental source: strain O157:H7, substrain EDL933  
C;Genetics:  
A;Gene: fimI  
C;Superfamily: type 1 fimbrial protein

Query Match 1.5%; Score 8; DB 2; Length 215;  
Best Local Similarity 100.0%; Pred. No. 6.5;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 LFLLASLL 11  
|||||||  
Db 42 LFLLASLL 49

#### RESULT 15

B91288  
fimbrial protein [imported] - Escherichia coli (strain O157:H7, substrain RIMD 0509952)  
C;Species: Escherichia coli  
C;Date: 18-Jul-2001 #sequence\_revision 18-Jul-2001 #text\_change 02-Aug-2002  
C;Accession: B91288  
R;Hayashi, T.; Makino, K.; Ohnishi, M.; Kurokawa, K.; Ishii, K.; Yokoyama, K.; Han, C.G.  
gasawara, N.; Yasunaga, T.; Kuhara, S.; Shiba, T.; Hattori, M.; Shinagawa, H.  
DNA Res. 8, 11-22, 2001  
A;Title: Complete genome sequence of enterohemorrhagic Escherichia coli O157:H7 and gene  
A;Reference number: A99629; MUID:21156231; PMID:11258796  
A;Accession: B91288  
A;Status: preliminary  
A;Molecule type: DNA  
A;Residues: 1-215 <HAY>  
A;Cross-references: GB:BA000007; PIDN:BA838697.1; PID:gl3364752; GSPDB:GN00154  
A;Experimental source: strain O157:H7, substrain RIMD 0509952  
C;Genetics:  
A;Gene: EC85274  
C;Superfamily: type 1 fimbrial protein

Query Match 1.5%; Score 8; DB 2; Length 215;  
Best Local Similarity 100.0%; Pred. No. 6.5;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 LFLLASLL 11  
|||||||  
Db 42 LFLLASLL 49

Search completed: May 5, 2004, 13:48:28  
Job time : 16.595 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 5, 2004, 13:33:49 ; Search time 42.2418 Seconds  
(without alignments)

3438.471 Million cell updates/sec

Title: US-09-743-684A-1

Perfect score: 524

Sequence: 1 MYALFLIASLLGAALAGPVL.....NTETAQCNAVERCKRHVWN 524

Scoring table: OLIGO

Gapop 60.0 , Gapext 60.0

Searched: 1138120 seqs, 277189581 residues

Word size : 0

Total number of hits satisfying chosen parameters: 1138120

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database :

- Published Applications AA:\*
- 1: /cgn2\_6/ptodata/2/pubpaa/US07\_PUBCOMB.pep.\*
  - 2: /cgn2\_6/ptodata/2/pubpaa/PCT\_NEW\_PUB.pep.\*
  - 3: /cgn2\_6/ptodata/2/pubpaa/US06\_NEW\_PUB.pep.\*
  - 4: /cgn2\_6/ptodata/2/pubpaa/US06\_PUBCOMB.pep.\*
  - 5: /cgn2\_6/ptodata/2/pubpaa/US07\_NEW\_PUB.pep.\*
  - 6: /cgn2\_6/ptodata/2/pubpaa/PCTUS\_PUBCOMB.pep.\*
  - 7: /cgn2\_6/ptodata/2/pubpaa/US08\_NEW\_PUB.pep.\*
  - 8: /cgn2\_6/ptodata/2/pubpaa/US08\_PUBCOMB.pep.\*
  - 9: /cgn2\_6/ptodata/2/pubpaa/US09A\_PUBCOMB.pep.\*
  - 10: /cgn2\_6/ptodata/2/pubpaa/US09B\_PUBCOMB.pep.\*
  - 11: /cgn2\_6/ptodata/2/pubpaa/US09C\_PUBCOMB.pep.\*
  - 12: /cgn2\_6/ptodata/2/pubpaa/US09\_NEW\_PUB.pep.\*
  - 13: /cgn2\_6/ptodata/2/pubpaa/US10A\_PUBCOMB.pep.\*
  - 14: /cgn2\_6/ptodata/2/pubpaa/US10B\_PUBCOMB.pep.\*
  - 15: /cgn2\_6/ptodata/2/pubpaa/US10C\_PUBCOMB.pep.\*
  - 16: /cgn2\_6/ptodata/2/pubpaa/US10\_NEW\_PUB.pep.\*
  - 17: /cgn2\_6/ptodata/2/pubpaa/US60\_NEW\_PUB.pep.\*
  - 18: /cgn2\_6/ptodata/2/pubpaa/US60\_PUBCOMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	524	100.0	524	9	US-09-870-759-60
2	524	100.0	524	10	US-09-751-708A-60
3	524	100.0	524	12	US-10-267-502-386
4	284	54.2	527	9	US-09-870-759-61
5	284	54.2	527	10	US-09-751-708A-61
6	284	54.2	527	14	US-10-060-036-73
7	260	49.6	479	10	US-09-978-418-40
8	209	39.9	209	14	US-10-043-487-340
9	193	36.8	523	9	US-09-767-007A-2
10	80	15.3	80	9	US-09-767-007A-3
11	80	15.3	80	9	US-09-753-126-3
12	80	15.3	80	15	US-10-330-697-3
13	80	15.3	592	9	US-09-753-126-4
14	80	15.3	592	15	US-10-330-697-4
15	40	7.6	40	10	US-09-780-438A-1

16	38	7.3	38	10	US-09-780-438A-2	Sequence 2, Appli
17	24	4.6	156	12	US-09-925-298-644	Sequence 644, App
18	156	4.6	156	14	US-10-102-806-644	Sequence 644, App
19	22	4.2	22	9	US-09-767-007A-1	Sequence 1, Appli
20	22	4.2	22	9	US-09-767-007A-11	Sequence 11, Appli
21	22	4.2	22	9	US-09-957-143-1	Sequence 1, Appli
22	22	4.2	22	9	US-09-957-143-1	Sequence 176, App
23	20	3.8	554	14	US-10-205-194-176	Sequence 5, Appli
24	17	3.2	18	9	US-09-767-007A-5	Sequence 2, Appli
25	17	3.2	18	9	US-09-957-143-2	Sequence 2, Appli
26	15	2.9	15	9	US-09-802-617-2	Sequence 2, Appli
27	15	2.9	15	14	US-10-293-819-2	Sequence 1, Appli
28	12	2.3	12	9	US-09-802-617-1	Sequence 3, Appli
29	12	2.3	12	9	US-09-957-143-3	Sequence 1, Appli
30	12	2.3	12	14	US-10-293-819-1	Sequence 182, App
31	11	2.1	11	10	US-09-791-393-182	Sequence 182, App
32	11	2.1	11	10	US-09-791-389-182	Sequence 126, App
33	10	1.9	15	14	US-10-281-478-126	Sequence 74, Appli
34	10	1.9	20	14	US-10-281-478-74	Sequence 75, Appli
35	9	1.7	14	14	US-10-281-478-75	Sequence 129, App
36	9	1.7	17	14	US-10-281-478-129	Sequence 362, App
37	8	1.5	215	9	US-09-912-020-362	Sequence 337, App
38	8	1.5	269	12	US-10-243-552-937	Sequence 4623, Ap
39	8	1.5	286	15	US-10-108-260A-4623	Sequence 279814,
40	8	1.5	354	12	US-10-424-599-279814	Sequence 46329, A
41	8	1.5	361	12	US-10-425-114-46329	Sequence 72368, A
42	8	1.5	363	12	US-10-425-114-72368	Sequence 2324, Ap
43	8	1.5	391	14	US-10-017-161-2324	Sequence 1970, Ap
44	8	1.5	391	15	US-10-292-798-1970	Sequence 82, Appli
45	8	1.5	521	12	US-10-412-699B-82	

ALIGNMENTS

RESULT 1

US-09-870-759-60  
; Sequence 60, Application US/09870759  
; Patent No. US20020177551A1  
; GENERAL INFORMATION:  
; APPLICANT: TERMAN, David S  
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATMENT OF NEOPLASTIC DISEASE  
; FILE REFERENCE: 870759  
; CURRENT APPLICATION NUMBER: US/09/870,759  
; CURRENT FILING DATE: 2002-01-14  
; PRIOR APPLICATION NUMBER: US 60/208,128  
; PRIOR FILING DATE: 2000-05-30  
; NUMBER OF SEQ ID NOS: 166  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 60  
; LENGTH: 524  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-870-759-60

Query Match	100.0%;	Score 524;	DB 9;	Length 524;
Best Local Similarity	100.0%;	Pred. No. 0;		
Matches 524;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
Qy	1	MYALFLIASLLGAALAGPVLGLKCEKTRGSVWCONVKTASDCGAVKHCLQTVWNKPTVKS	60	
Db	1	MYALFLIASLLGAALAGPVLGLKCEKTRGSVWCONVKTASDCGAVKHCLQTVWNKPTVKS	60	
Qy	61	LPDCICKDVVTAAGDMLKDNATEEIIIVYLEKTCMDLKPKNMSASCKEIVDSYLPVILDI	120	
Db	61	LPDCICKDVVTAAGDMLKDNATEEIIIVYLEKTCMDLKPKNMSASCKEIVDSYLPVILDI	120	
Qy	121	IKGEMSRPGVCSALNLCESLQKHAEINHQKLESNKIPELDMTEVVAPEMANIPILLY	180	
Db	121	IKGEMSRPGVCSALNLCESLQKHAEINHQKLESNKIPELDMTEVVAPEMANIPILLY	180	
Qy	181	PDGPRSKPQKONGDVYCDQCIQWVTDIQTAVRTNSTFVQALVEHVKECDRLGPGMADI	240	

Db 181 PDGPRSKPQKONGDVQCDCIQMVTDTQTAVRTNSTFVQALVEHVHKECDRLGPGMADI 240  
QY 241 CKNYISQYSEIATOMMMHMQPKIEICALVGFCDVEKMPMQTLVPAKASKNVI PALELVE 300  
Db 241 CKNYISQYSEIATOMMMHMQPKIEICALVGFCDVEKMPMQTLVPAKASKNVI PALELVE 300  
QY 301 PIKKEVPAKSDVYCEVCEFLVKEVTKLIDNNKTEKIEILDAFDKMCCKLPSLSEECQEV 360  
Db 301 PIKKEVPAKSDVYCEVCEFLVKEVTKLIDNNKTEKIEILDAFDKMCCKLPSLSEECQEV 360  
QY 361 VDTYSSILSILLEVSPBELVCSMLHLCSTGTRLPALTTHVTQPKDGGFCEVCKKLVGYLD 420  
Db 361 VDTYSSILSILLEVSPBELVCSMLHLCSTGTRLPALTTHVTQPKDGGFCEVCKKLVGYLD 420  
QY 421 RNLEKNTKQEIILAALEKGCFLPDYQKQCDQFVAEYEPVLIEILVEVMDPSFVCLKIG 480  
Db 421 RNLEKNTKQEIILAALEKGCFLPDYQKQCDQFVAEYEPVLIEILVEVMDPSFVCLKIG 480  
QY 481 ACPSAHKPLLGTEKCIWGPSYWCNTETAACNAVEHCKRHVN 524  
Db 481 ACPSAHKPLLGTEKCIWGPSYWCNTETAACNAVEHCKRHVN 524

## RESULT 2

US-09-751-708A-60  
; Sequence 60, Application US/09751708A  
; Publication No. US20030157113A1  
; GENERAL INFORMATION:  
; APPLICANT: TERMAN, David S  
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATMENT OF NEOPLASTIC DISEASE  
; FILE REFERENCE: 751708  
; CURRENT APPLICATION NUMBER: US/09/751,708A  
; CURRENT FILING DATE: 2002-10-15  
; PRIOR APPLICATION NUMBER: US 60/173,371  
; FILING DATE: 1999-12-28  
; NUMBER OF SEQ ID NOS: 166  
; SOFTWARE: Patent in version 3.1  
; SEQ ID NO 60  
; LENGTH: 524  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-751-708A-60

Query Match 100.0%; Score 524; DB 10; Length 524;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 524; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 MYALFLASLLGAALAGPVLGLKECTRGSAVWCQNVKTASDCGAVKHCLQTVWKNKPTVKS 60  
Db 1 MYALFLASLLGAALAGPVLGLKECTRGSAVWCQNVKTASDCGAVKHCLQTVWKNKPTVKS 60  
QY 61 LPCDICDVVTAAGDMLKDNATBEEILVYLEKTCDWLPKPNMSASCKEIVDSYLPVILDI 120  
Db 61 LPCDICDVVTAAGDMLKDNATBEEILVYLEKTCDWLPKPNMSASCKEIVDSYLPVILDI 120  
QY 121 IKGEMSPGVCESALNLCESLQKHLAELNHOKLESNKIPELDMTEVAPFMANIPLLLY 180  
Db 121 IKGEMSPGVCESALNLCESLQKHLAELNHOKLESNKIPELDMTEVAPFMANIPLLLY 180  
QY 181 PDGPRSKPQKONGDVQCDCIQMVTDTQTAVRTNSTFVQALVEHVHKECDRLGPGMADI 240  
Db 181 PDGPRSKPQKONGDVQCDCIQMVTDTQTAVRTNSTFVQALVEHVHKECDRLGPGMADI 240  
QY 241 CKNYISQYSEIATOMMMHMQPKIEICALVGFCDVEKMPMQTLVPAKASKNVI PALELVE 300  
Db 241 CKNYISQYSEIATOMMMHMQPKIEICALVGFCDVEKMPMQTLVPAKASKNVI PALELVE 300  
QY 301 PIKKEVPAKSDVYCEVCEFLVKEVTKLIDNNKTEKIEILDAFDKMCCKLPSLSEECQEV 360  
Db 301 PIKKEVPAKSDVYCEVCEFLVKEVTKLIDNNKTEKIEILDAFDKMCCKLPSLSEECQEV 360  
QY 361 VDTYSSILSILLEVSPBELVCSMLHLCSTGTRLPALTTHVTQPKDGGFCEVCKKLVGYLD 420

Db 361 VDTYSSILSILLEVSPBELVCSMLHLCSTGTRLPALTTHVTQPKDGGFCEVCKKLVGYLD 420  
QY 421 RNLEKNTKQEIILAALEKGCFLPDYQKQCDQFVAEYEPVLIEILVEVMDPSFVCLKIG 480  
Db 421 RNLEKNTKQEIILAALEKGCFLPDYQKQCDQFVAEYEPVLIEILVEVMDPSFVCLKIG 480  
QY 481 ACPSAHKPLLGTEKCIWGPSYWCNTETAACNAVEHCKRHVN 524  
Db 481 ACPSAHKPLLGTEKCIWGPSYWCNTETAACNAVEHCKRHVN 524

## RESULT 3

US-10-267-502-386  
; Sequence 386, Application US/10267502  
; Publication No. US20040071700A1  
; GENERAL INFORMATION:  
; APPLICANT: Kim, Jaeseob  
; APPLICANT: Galant, Ron  
; TITLE OF INVENTION: Obesity Linked Genes  
; FILE REFERENCE: LSD-07416  
; CURRENT APPLICATION NUMBER: US/10/267,502  
; CURRENT FILING DATE: 2003-01-27  
; NUMBER OF SEQ ID NOS: 439  
; SOFTWARE: Patent in version 3.2  
; SEQ ID NO 386  
; LENGTH: 524  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-267-502-386

Query Match 100.0%; Score 524; DB 12; Length 524;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 524; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 MYALFLASLLGAALAGPVLGLKECTRGSAVWCQNVKTASDCGAVKHCLQTVWKNKPTVKS 60  
Db 1 MYALFLASLLGAALAGPVLGLKECTRGSAVWCQNVKTASDCGAVKHCLQTVWKNKPTVKS 60  
QY 61 LPCDICDVVTAAGDMLKDNATBEEILVYLEKTCDWLPKPNMSASCKEIVDSYLPVILDI 120  
Db 61 LPCDICDVVTAAGDMLKDNATBEEILVYLEKTCDWLPKPNMSASCKEIVDSYLPVILDI 120  
QY 121 IKGEMSPGVCESALNLCESLQKHLAELNHOKLESNKIPELDMTEVAPFMANIPLLLY 180  
Db 121 IKGEMSPGVCESALNLCESLQKHLAELNHOKLESNKIPELDMTEVAPFMANIPLLLY 180  
QY 181 PDGPRSKPQKONGDVQCDCIQMVTDTQTAVRTNSTFVQALVEHVHKECDRLGPGMADI 240  
Db 181 PDGPRSKPQKONGDVQCDCIQMVTDTQTAVRTNSTFVQALVEHVHKECDRLGPGMADI 240  
QY 241 CKNYISQYSEIATOMMMHMQPKIEICALVGFCDVEKMPMQTLVPAKASKNVI PALELVE 300  
Db 241 CKNYISQYSEIATOMMMHMQPKIEICALVGFCDVEKMPMQTLVPAKASKNVI PALELVE 300  
QY 301 PIKKEVPAKSDVYCEVCEFLVKEVTKLIDNNKTEKIEILDAFDKMCCKLPSLSEECQEV 360  
Db 301 PIKKEVPAKSDVYCEVCEFLVKEVTKLIDNNKTEKIEILDAFDKMCCKLPSLSEECQEV 360  
QY 361 VDTYSSILSILLEVSPBELVCSMLHLCSTGTRLPALTTHVTQPKDGGFCEVCKKLVGYLD 420  
Db 361 VDTYSSILSILLEVSPBELVCSMLHLCSTGTRLPALTTHVTQPKDGGFCEVCKKLVGYLD 420  
QY 421 RNLEKNTKQEIILAALEKGCFLPDYQKQCDQFVAEYEPVLIEILVEVMDPSFVCLKIG 480  
Db 421 RNLEKNTKQEIILAALEKGCFLPDYQKQCDQFVAEYEPVLIEILVEVMDPSFVCLKIG 480  
QY 481 ACPSAHKPLLGTEKCIWGPSYWCNTETAACNAVEHCKRHVN 524  
Db 481 ACPSAHKPLLGTEKCIWGPSYWCNTETAACNAVEHCKRHVN 524

## RESULT 4

US-09-870-759-61

```
; Sequence 61, Application US/09870759
; Patent No. US20020177551A1
; GENERAL INFORMATION:
; APPLICANT: TERMAN, David S
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATMENT OF NEOPLASTIC DISEASE
; FILE REFERENCE: 870759
; CURRENT APPLICATION NUMBER: US/09/870,759
; CURRENT FILING DATE: 2002-01-14
; PRIOR APPLICATION NUMBER: US 60/208,128
; PRIOR FILING DATE: 2000-05-30
; NUMBER OF SEQ ID NOS: 166
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 61
; TYPE: PR1
; ORGANISM: Homo sapiens
US-09-870-759-61

Query Match      54.2%; Score 284; DB 9; Length 527;
Best Local Similarity 99.4%; Pred. No. 1.1e-267;
Matches 524; Conservative 0; Mismatches 0; Indels 3; Gaps 1;

Qy 1 MYALFLLASLLGAALAGPVLGLKECTRGSAVWCNQVKTASDCGAVKHCLQTVWKNKPTVKS 60
Db 1 MYALFLLASLLGAALAGPVLGLKECTRGSAVWCNQVKTASDCGAVKHCLQTVWKNKPTVKS 60
Qy 61 LPCDICKDVVTAAGDMLKDNATEEEILVYLEKTCDWLPKPNMSASCKEIVDSYLPVILDI 120
Db 61 LPCDICKDVVTAAGDMLKDNATEEEILVYLEKTCDWLPKPNMSASCKEIVDSYLPVILDI 120
Qy 121 IKGMSRPGVCSALNLCESLQKHLAELNHQKLESNKIPELDMTEVVAFFMANIPLLLY 180
Db 121 IKGMSRPGVCSALNLCESLQKHLAELNHQKLESNKIPELDMTEVVAFFMANIPLLLY 180
Qy 181 PQDGRSKPOPKONGDVQCDCIQMWTDIQTAVRTNSTFVQALVEHVKEECDRLGPGMADI 240
Db 181 PQDGRSKPOPKONGDVQCDCIQMWTDIQTAVRTNSTFVQALVEHVKEECDRLGPGMADI 240
Qy 241 CKNYISOYSEIAIOMMHM--OPKEICALVGFCDVEKMPQTLVPAKVASKNVIIPALE 297
Db 241 CKNYISOYSEIAIOMMHMDDQPKETCALVGFCDVEKMPQTLVPAKVASKNVIIPALE 300
Qy 298 LVEPIKKEHVPKSDVYCEVCEFLVKEVTKLIDNNKTEKEILDADFDMCKSLPKSLSEEC 357
Db 301 LVEPIKKEHVPKSDVYCEVCEFLVKEVTKLIDNNKTEKEILDADFDMCKSLPKSLSEEC 360
Qy 358 QEVVDYTGSSILSLLEEVSPVLCVSMHLCSGTRLPALTVHVTQPKDGGFCEVCKKLVG 417
Db 361 QEVVDYTGSSILSLLEEVSPVLCVSMHLCSGTRLPALTVHVTQPKDGGFCEVCKKLVG 420
Qy 418 YLDRNLEKNSTKQELIAALEKGCFLPDPYKQKCDQFVAEYEPVLIILVEVMDPSFVCL 477
Db 421 YLDRNLEKNSTKQELIAALEKGCFLPDPYKQKCDQFVAEYEPVLIILVEVMDPSFVCL 480
Qy 478 KIGACPSAHKPLLGTEKRCINGSPSYWCQNTETAACNAVEHCKRHVWN 524
Db 481 KIGACPSAHKPLLGTEKRCINGSPSYWCQNTETAACNAVEHCKRHVWN 527

RESULT 5
US-09-751-708A-61
; Sequence 61, Application US/09751708A
; Publication No. US20030157113A1
; GENERAL INFORMATION:
; APPLICANT: TERMAN, David S
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATMENT OF NEOPLASTIC DISEASE
; FILE REFERENCE: 751708
; CURRENT APPLICATION NUMBER: US/09/751,708A
; CURRENT FILING DATE: 2002-10-15
; PRIOR APPLICATION NUMBER: US 60/173,371
; PRIOR FILING DATE: 1999-12-28
; NUMBER OF SEQ ID NOS: 166
; SOFTWARE: PatentIn version 3.1
```

```
; SEQ ID NO 61
; LENGTH: 527
; TYPE: PR1
; ORGANISM: Homo sapiens
US-09-751-708A-61

Query Match      54.2%; Score 284; DB 10; Length 527;
Best Local Similarity 99.4%; Pred. No. 1.1e-267;
Matches 524; Conservative 0; Mismatches 0; Indels 3; Gaps 1;

Qy 1 MYALFLLASLLGAALAGPVLGLKECTRGSAVWCNQVKTASDCGAVKHCLQTVWKNKPTVKS 60
Db 1 MYALFLLASLLGAALAGPVLGLKECTRGSAVWCNQVKTASDCGAVKHCLQTVWKNKPTVKS 60
Qy 61 LPCDICKDVVTAAGDMLKDNATEEEILVYLEKTCDWLPKPNMSASCKEIVDSYLPVILDI 120
Db 61 LPCDICKDVVTAAGDMLKDNATEEEILVYLEKTCDWLPKPNMSASCKEIVDSYLPVILDI 120
Qy 121 IKGMSRPGVCSALNLCESLQKHLAELNHQKLESNKIPELDMTEVVAFFMANIPLLLY 180
Db 121 IKGMSRPGVCSALNLCESLQKHLAELNHQKLESNKIPELDMTEVVAFFMANIPLLLY 180
Qy 181 PQDGRSKPOPKONGDVQCDCIQMWTDIQTAVRTNSTFVQALVEHVKEECDRLGPGMADI 240
Db 181 PQDGRSKPOPKONGDVQCDCIQMWTDIQTAVRTNSTFVQALVEHVKEECDRLGPGMADI 240
Qy 241 CKNYISOYSEIAIOMMHM--OPKEICALVGFCDVEKMPQTLVPAKVASKNVIIPALE 297
Db 241 CKNYISOYSEIAIOMMHMDDQPKETCALVGFCDVEKMPQTLVPAKVASKNVIIPALE 300
Qy 298 LVEPIKKEHVPKSDVYCEVCEFLVKEVTKLIDNNKTEKEILDADFDMCKSLPKSLSEEC 357
Db 301 LVEPIKKEHVPKSDVYCEVCEFLVKEVTKLIDNNKTEKEILDADFDMCKSLPKSLSEEC 360
Qy 358 QEVVDYTGSSILSLLEEVSPVLCVSMHLCSGTRLPALTVHVTQPKDGGFCEVCKKLVG 417
Db 361 QEVVDYTGSSILSLLEEVSPVLCVSMHLCSGTRLPALTVHVTQPKDGGFCEVCKKLVG 420
Qy 418 YLDRNLEKNSTKQELIAALEKGCFLPDPYKQKCDQFVAEYEPVLIILVEVMDPSFVCL 477
Db 421 YLDRNLEKNSTKQELIAALEKGCFLPDPYKQKCDQFVAEYEPVLIILVEVMDPSFVCL 480
Qy 478 KIGACPSAHKPLLGTEKRCINGSPSYWCQNTETAACNAVEHCKRHVWN 524
Db 481 KIGACPSAHKPLLGTEKRCINGSPSYWCQNTETAACNAVEHCKRHVWN 527

RESULT 6
US-10-060-036-73
; Sequence 73, Application US/10060036
; Publication No. US20030073144A1
; GENERAL INFORMATION:
; APPLICANT: Benson, Darin R.
; APPLICANT: Kalos, Michael D.
; APPLICANT: Lodes, Michael J.
; APPLICANT: Persing, David H.
; APPLICANT: Hepler, William T.
; APPLICANT: Jiang, Yuqiu
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; FILE REFERENCE: 210121.566
; CURRENT APPLICATION NUMBER: US/10/060,036
; CURRENT FILING DATE: 2002-01-30
; NUMBER OF SEQ ID NOS: 4560
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 73
; LENGTH: 527
; TYPE: PR1
; ORGANISM: Homo sapiens
US-10-060-036-73

Query Match      54.2%; Score 284; DB 14; Length 527;
Best Local Similarity 99.4%; Pred. No. 1.1e-267;
```

Matches 524; Conservative 0; Mismatches 0; Indels 3; Gaps 1;

QY 1 MYALFLLASLLGALAGPVLGLKECTRGSAVWQNVKTASDCGAVKHCLQTWNKPTVKS 60  
Db 1 MYALFLLASLLGALAGPVLGLKECTRGSAVWQNVKTASDCGAVKHCLQTWNKPTVKS 60

QY 61 LPCDICKDVVTAAGDMLKDNATEEEILVYLEKTCMDLKPNNMSASCKEIVDSYLPVILDI 120  
Db 61 LPCDICKDVVTAAGDMLKDNATEEEILVYLEKTCMDLKPNNMSASCKEIVDSYLPVILDI 120

QY 121 IKGEMSRPGEVCSALNLCESLQKHLAELNHQKLESNKIPELDMTEVWAPFMANIPLLLY 180  
Db 121 IKGEMSRPGEVCSALNLCESLQKHLAELNHQKLESNKIPELDMTEVWAPFMANIPLLLY 180

QY 181 PQDGRSKPQKNDGVDVCDICQMVTDIQTAVRTNSTFVQALVEHVKEECDRLGPGMADI 240  
Db 181 PQDGRSKPQKNDGVDVCDICQMVTDIQTAVRTNSTFVQALVEHVKEECDRLGPGMADI 240

QY 241 CKNYISQYSEIAIQMMHMQ 260  
Db 241 CKNYISQYSEIAIQMMHMQ 260

RESULT 8  
US-10-043-487-340  
; Sequence 340, Application US/10043487  
; Publication No. US2003005220A1  
; GENERAL INFORMATION:  
; APPLICANT: HYBRIGENICS  
; APPLICANT: Pierre, LEGRAN  
; TITLE OF INVENTION: Protein-protein interactions between Shigella Flexneri polypeptides  
; FILE REFERENCE: B4778A  
; CURRENT APPLICATION NUMBER: US/10/043,487  
; CURRENT FILING DATE: 2002-04-30  
; PRIOR APPLICATION NUMBER: US 60/261,130  
; PRIOR FILING DATE: 2001-01-12  
; NUMBER OF SEQ ID NOS: 561  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 340  
; LENGTH: 209  
; TYPE: PRT  
; ORGANISM: Shigella Flexneri  
US-10-043-487-340

Query Match 39.9%; Score 209; DB 14; Length 209;  
Best Local Similarity 100.0%; Pred. No. 5.7e-195;  
Matches 209; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 151 OKLESNKIPELDMTEVWAPFMANIPLLYPQDGRSKPQKNDGVDVCDICQMVTDIQT 210  
Db 1 OKLESNKIPELDMTEVWAPFMANIPLLYPQDGRSKPQKNDGVDVCDICQMVTDIQT 60

QY 211 AVRTNSTFVQALVEHVKEECDRLGPGMADICKNYISOYSEIAIQMMHMQKEICLVGF 270  
Db 61 AVRTNSTFVQALVEHVKEECDRLGPGMADICKNYISOYSEIAIQMMHMQKEICLVGF 120

QY 271 CDEVKEMPQTLVPAKVASKNVIPALELVEPIKKEHVPKASDVYCEVCEFLVKEVTKLID 330  
Db 121 CDEVKEMPQTLVPAKVASKNVIPALELVEPIKKEHVPKASDVYCEVCEFLVKEVTKLID 180

QY 331 NNKTEKILDAFDKMSKLPKSLSEECQE 359  
Db 181 NNKTEKILDAFDKMSKLPKSLSEECQE 209

RESULT 9  
US-09-767-007A-2  
; Sequence 2, Application US/09767007A  
; Patent No. US20020077275A1  
; GENERAL INFORMATION:  
; APPLICANT: John S. O'Brien  
; APPLICANT: Yasuo Kishimoto  
; TITLE OF INVENTION: PHARMACEUTICAL COMPOSITIONS COMPRISING  
; FILE REFERENCE: SAPOSIN C AND NEUROTROPIC PEPTIDES DERIVED THEREFROM  
; CURRENT APPLICATION NUMBER: US/09/767,007A  
; CURRENT FILING DATE: 2001-01-22

Matches 524; Conservative 0; Mismatches 0; Indels 3; Gaps 1;

QY 1 MYALFLLASLLGALAGPVLGLKECTRGSAVWQNVKTASDCGAVKHCLQTWNKPTVKS 60  
Db 1 MYALFLLASLLGALAGPVLGLKECTRGSAVWQNVKTASDCGAVKHCLQTWNKPTVKS 60

QY 61 LPCDICKDVVTAAGDMLKDNATEEEILVYLEKTCMDLKPNNMSASCKEIVDSYLPVILDI 120  
Db 61 LPCDICKDVVTAAGDMLKDNATEEEILVYLEKTCMDLKPNNMSASCKEIVDSYLPVILDI 120

QY 121 IKGEMSRPGEVCSALNLCESLQKHLAELNHQKLESNKIPELDMTEVWAPFMANIPLLLY 180  
Db 121 IKGEMSRPGEVCSALNLCESLQKHLAELNHQKLESNKIPELDMTEVWAPFMANIPLLLY 180

QY 181 PQDGRSKPQKNDGVDVCDICQMVTDIQTAVRTNSTFVQALVEHVKEECDRLGPGMADI 240  
Db 181 PQDGRSKPQKNDGVDVCDICQMVTDIQTAVRTNSTFVQALVEHVKEECDRLGPGMADI 240

QY 241 CKNYISQYSEIAIQMMHMQ 260  
Db 241 CKNYISQYSEIAIQMMHMQ 260

RESULT 8  
US-10-043-487-340  
; Sequence 340, Application US/10043487  
; Publication No. US2003005220A1  
; GENERAL INFORMATION:  
; APPLICANT: HYBRIGENICS  
; APPLICANT: Pierre, LEGRAN  
; TITLE OF INVENTION: Protein-protein interactions between Shigella Flexneri polypeptides  
; FILE REFERENCE: B4778A  
; CURRENT APPLICATION NUMBER: US/10/043,487  
; CURRENT FILING DATE: 2002-04-30  
; PRIOR APPLICATION NUMBER: US 60/261,130  
; PRIOR FILING DATE: 2001-01-12  
; NUMBER OF SEQ ID NOS: 561  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 340  
; LENGTH: 209  
; TYPE: PRT  
; ORGANISM: Shigella Flexneri  
US-10-043-487-340

```
; PRIOR APPLICATION NUMBER: 08/958,970
; PRIOR FILING DATE: 1997-10-28
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 523
; TYPE: PRT
; ORGANISM: homo sapiens
US-09-767-007A-2

Query Match          36.8%; Score 193; DB 9; Length 523;
Best Local Similarity 100.0%; Pred. No. 5.1e-179; Indels 0; Gaps 0;
Matches 193; Conservative 0; Mismatches 0;

Qy 1 MYALFLASLLGAALAGPVLGLKECTRGSAVWCONVKTASDCGAVKHCLQTVWNKPTVKS 60
    |||||||
Db 1 MYALFLASLLGAALAGPVLGLKECTRGSAVWCONVKTASDCGAVKHCLQTVWNKPTVKS 60
    |||||||

Qy 61 LPDCICKDVVTAAGDMLKDNATEEILVYLEKTCDWLPKPNMGASCKEIVDSYLPVILDI 120
    |||||||
Db 61 LPDCICKDVVTAAGDMLKDNATEEILVYLEKTCDWLPKPNMGASCKEIVDSYLPVILDI 120
    |||||||

Qy 121 IKGEMSPGVCSALNLCESIQKHLAELNHQKLESNKIPELDMTEVVAPFMANIPILLY 180
    |||||||
Db 121 IKGEMSPGVCSALNLCESIQKHLAELNHQKLESNKIPELDMTEVVAPFMANIPILLY 180
    |||||||

Qy 181 PQDGPSPKPOPKD 193
    |||||||
Db 181 PQDGPSPKPOPKD 193
    |||||||

RESULT 10
US-09-767-007A-3
; Sequence 3, Application US/09767007A
; Patent No. US2002007275A1
; GENERAL INFORMATION:
; APPLICANT: John S. O'Brien
; APPLICANT: Yasuo Kishimoto
; TITLE OF INVENTION: PHARMACEUTICAL COMPOSITIONS COMPRISING
; FILE OF INVENTION: SAPOSIN C AND NEUROTROPIC PEPTIDES DERIVED THEREFROM
; CURRENT APPLICATION NUMBER: US/09/767,007A
; CURRENT FILING DATE: 2001-01-22
; PRIOR APPLICATION NUMBER: 08/958,970
; PRIOR FILING DATE: 1997-10-28
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 80
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Artificial peptide
US-09-767-007A-3

Query Match          15.3%; Score 80; DB 9; Length 80;
Best Local Similarity 100.0%; Pred. No. 1.2e-69; Indels 0; Gaps 0;
Matches 80; Conservative 0; Mismatches 0;

Qy 311 SDVYCEVCEFLVKEVTKLIDNNKTEKEILDADFDMCKSLPKSLSEECQEVVDYTGSSILS 370
    |||||||
Db 1 SDVYCEVCEFLVKEVTKLIDNNKTEKEILDADFDMCKSLPKSLSEECQEVVDYTGSSILS 60
    |||||||

Qy 371 ILLEEVSPPELVCSMLHLCSG 390
    |||||||
Db 61 ILLEEVSPPELVCSMLHLCSG 80
    |||||||

RESULT 11
US-09-753-126-3
; Sequence 3, Application US/09753126
; Patent No. US2002012719A1
; GENERAL INFORMATION:
```

```
; APPLICANT: OKKELS, JENS SIGURD
; APPLICANT: JENSEN, ANNE DAM
; APPLICANT: HALKIER, TORBEN
; APPLICANT: JENSEN, RIKKE BOLDING
; TITLE OF INVENTION: IMPROVED LYSOSOMAL ENZYMES AND LYSOSOMAL ENZYME
; TITLE OF INVENTION: ACTIVATORS
; FILE REFERENCE: 31-0006000S
; CURRENT APPLICATION NUMBER: US/09/753,126
; CURRENT FILING DATE: 2001-06-11
; PRIOR APPLICATION NUMBER: PA 1999 01891
; PRIOR FILING DATE: 1999-12-30
; PRIOR APPLICATION NUMBER: 60/174,652
; PRIOR FILING DATE: 2000-01-06
; PRIOR APPLICATION NUMBER: PA 200 00865
; PRIOR FILING DATE: 2000-06-02
; PRIOR APPLICATION NUMBER: 60/210,984
; PRIOR FILING DATE: 2000-06-12
; PRIOR APPLICATION NUMBER: 60/211,124
; PRIOR FILING DATE: 2000-06-12
; PRIOR APPLICATION NUMBER: PA 2000 01027
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: 60/217,497
; PRIOR FILING DATE: 2000-07-11
; NUMBER OF SEQ ID NOS: 147
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 3
; LENGTH: 80
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-753-126-3

Query Match          15.3%; Score 80; DB 9; Length 80;
Best Local Similarity 100.0%; Pred. No. 1.2e-69; Indels 0; Gaps 0;
Matches 80; Conservative 0; Mismatches 0;

Qy 311 SDVYCEVCEFLVKEVTKLIDNNKTEKEILDADFDMCKSLPKSLSEECQEVVDYTGSSILS 370
    |||||||
Db 1 SDVYCEVCEFLVKEVTKLIDNNKTEKEILDADFDMCKSLPKSLSEECQEVVDYTGSSILS 60
    |||||||

Qy 371 ILLEEVSPPELVCSMLHLCSG 390
    |||||||
Db 61 ILLEEVSPPELVCSMLHLCSG 80
    |||||||

RESULT 12
US-10-330-697-3
; Sequence 3, Application US/10330697
; Publication No. US20040009165A1
; GENERAL INFORMATION:
; APPLICANT: OKKELS, JENS SIGURD
; APPLICANT: JENSEN, ANNE DAM
; APPLICANT: HALKIER, TORBEN
; APPLICANT: JENSEN, RIKKE BOLDING
; TITLE OF INVENTION: IMPROVED LYSOSOMAL ENZYMES AND LYSOSOMAL ENZYME
; TITLE OF INVENTION: ACTIVATORS
; FILE REFERENCE: 31-0006000S
; CURRENT APPLICATION NUMBER: US/10/330,697
; CURRENT FILING DATE: 2002-12-27
; PRIOR APPLICATION NUMBER: US/09/753,126
; PRIOR FILING DATE: 2001-06-11
; PRIOR APPLICATION NUMBER: PA 1999 01891
; PRIOR FILING DATE: 1999-12-30
; PRIOR APPLICATION NUMBER: 60/174,652
; PRIOR FILING DATE: 2000-01-06
; PRIOR APPLICATION NUMBER: PA 200 00865
; PRIOR FILING DATE: 2000-06-02
; PRIOR APPLICATION NUMBER: 60/210,984
; PRIOR FILING DATE: 2000-06-12
; PRIOR APPLICATION NUMBER: 60/211,124
; PRIOR FILING DATE: 2000-06-12
; PRIOR APPLICATION NUMBER: PA 2000 01027
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: 60/217,497
```

```

; PRIOR FILING DATE: 2000-07-11
; NUMBER OF SEQ ID NOS: 147
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 3
; LENGTH: 80
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-330-697-3

```

```
Query Match          15.3%; Score 80; DB 15; Length 80;
Best Local Similarity 100.0%; Pred. No. 1.2e-69;
Matches 80; Conservative 0; Mismatches 0; Indels
```

RESULT 13  
US-09-753-126-4  
; Sequence 4, Application US/09753126  
; Patent No. US20020127219A1  
; GENERAL INFORMATION:  
; APPLICANT: OKKELS, JENS SIGURD  
; APPLICANT: JENSEN, ANNE DAM  
; APPLICANT: HALKIER, TORBEN  
; APPLICANT: JENSEN, RIKKE BOLDING  
; TITLE OF INVENTION: IMPROVED LYSOSOMAL ENZYMES AND LYSOSOMAL ENZYME  
; TITLE OF INVENTION: ACTIVATORS  
; FILE REFERENCE: 31-0006000US

```

; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Chimeric
; OTHER INFORMATION: SapC-linker-GBB polypeptide
US-09-753-126-4

```

```

RESULT 14
US-10-330-697-4
; Sequence 4, Application US/10330697
; Publication No. US20040009165A1
; GENERAL INFORMATION:
; APPLICANT: OKKELS, JENS SIGURD
; APPLICANT: JENSEN, ANNE DAM
; APPLICANT: HALKIER, TORBEN
; APPLICANT: JENSEN, RIKKE BOLDING
; TITLE OF INVENTION: IMPROVED LYSOSOMAL ENZYMES AND LYSOSOMAL ENZYME
; TITLE OF INVENTION: ACTIVATORS
;

```

Query Match	15.3%	Score 80;	DB 15;	Length 592;
Best Local Similarity	100.0%;	Pred. No. 7, 1e-69;		
Matches 80;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
QY	311	SDVCEVCEFLVKEVTKLIDNNKTEKELIDAFDKMKSKLPKSLSECEQVVDVTYGSSTLS	370	
Db	1	SDVCEVCEFLVKEVTKLIDNNKTEKELIDAFDKMKSKLPKSLSECEQVVDVTYGSSTLS	60	
QY	371	ILLEEVSPELVCSMLHLCSG	390	
Db	61	ILLEEVSPELVCSMLHLCSG	80	

RESULT 15  
 US-09-780-438A-1  
 ; Sequence 1, Application US/09780438A  
 ; Publication No. US2003009599A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Cincinnati Childrens Hospital Research Foundation  
 ; TITLE OF INVENTION: Fusogenic Properties of Saposin C and Related Proteins and Polypeptides  
 ; TITLE OF INVENTION: for Application to Transmembrane Drug Delivery Systems  
 ; FILE REFERENCE: 10872/0474352  
 ; CURRENT APPLICATION NUMBER: US/09/780,438A  
 ; CURRENT FILING DATE: 2001-02-09  
 ; PRIOR APPLICATION NUMBER: US 60/181,754  
 ; PRIOR FILING DATE: 2000-02-11  
 ; NUMBER OF SEQ ID NOS: 2  
 ; SOFTWARE: PatentIn version 3.0  
 ; SEQ ID NO 1  
 ; LENGTH: 40  
 ; TYPE: PRT



; ORGANISM: Homo sapiens  
US-09-780-438A-1

Query Match 7.6%; Score 40; DB 10; Length 40;  
Best Local Similarity 100.0%; Pred. No. 5.7e-31;  
Matches 40; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 311 SDVYCEVCEFLVKEVTKLIDNNKTEKILDAFDKMCCKLP 350  
|||||  
Db 1 SDVYCEVCEFLVKEVTKLIDNNKTEKILDAFDKMCCKLP 40  
|||||

Search completed: May 5, 2004, 13:43:39  
Job time : 43.2418 secs

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 5, 2004, 13:28:28 ; Search time 16.595 Seconds  
(without alignments)  
1630.129 Million cell updates/sec

Title: US-09-743-684A-1

Perfect score: 524

Sequence: 1 MYALFLIASLGAALAGPVL.....NTETAACNAVERCKRHVWN 524

Scoring table: OLIGO

Gapop 60.0 , Gapext 60.0

Searched: 389414 seqs, 51625971 residues

Word size : 0

Total number of hits satisfying chosen parameters: 389414

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : Issued Patents AA.\*

- 1: /cgn2\_6/ptodata/2/iaa/5A COMB.pep.\*
- 2: /cgn2\_6/ptodata/2/iaa/5B COMB.pep.\*
- 3: /cgn2\_6/ptodata/2/iaa/6A COMB.pep.\*
- 4: /cgn2\_6/ptodata/2/iaa/6B COMB.pep.\*
- 5: /cgn2\_6/ptodata/2/iaa/PTCUS COMB.pep.\*
- 6: /cgn2\_6/ptodata/2/iaa/backfiles1.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	524	100.0	524	4	US-09-352-548-1
2	224	42.7	523	1	US-08-100-247-2
3	224	42.7	523	1	US-08-232-513A-3
4	224	42.7	523	4	US-08-756-031-2
5	193	36.8	523	1	US-08-483-146A-2
6	193	36.8	523	1	US-08-484-594A-2
7	193	36.8	523	4	US-09-076-258A-2
8	81	15.5	81	4	US-09-352-548-2
9	80	15.3	80	1	US-08-100-247-3
10	80	15.3	80	1	US-08-483-146A-3
11	80	15.3	80	1	US-08-232-513A-4
12	80	15.3	80	1	US-08-484-594A-3
13	80	15.3	80	4	US-09-076-258A-3
14	80	15.3	80	4	US-08-756-031-3
15	75	14.3	80	2	US-08-584-671-15
16	75	14.3	80	3	US-09-027-376-15
17	69	13.2	80	3	US-09-094-192-15
18	61	11.6	61	2	US-08-584-671-13
19	61	11.6	61	3	US-09-027-376-13
20	61	11.6	61	3	US-09-094-192-13
21	61	11.6	69	3	US-08-268-070-2
22	22	4.2	22	1	US-08-100-247-1
23	22	4.2	22	1	US-08-483-146A-1
24	22	4.2	22	1	US-08-483-146A-11
25	22	4.2	22	1	US-08-232-513A-1
26	22	4.2	22	1	US-08-232-513A-20
27	22	4.2	22	1	US-08-484-594A-1

28	22	4.2	22	1	US-08-484-594A-11	Sequence 11, Appl
29	22	4.2	22	3	US-09-231-159-1	Sequence 1, Appl
30	22	4.2	22	3	US-08-611-307-1	Sequence 1, Appl
31	22	4.2	22	4	US-09-148-030-1	Sequence 1, Appl
32	22	4.2	22	4	US-09-076-258A-1	Sequence 1, Appl
33	22	4.2	22	4	US-09-076-258A-11	Sequence 11, Appl
34	22	4.2	22	4	US-08-756-031-1	Sequence 1, Appl
35	18	3.4	18	4	US-09-352-548-17	Sequence 17, Appl
36	17	3.2	17	4	US-09-352-548-14	Sequence 14, Appl
37	17	3.2	17	4	US-09-352-548-15	Sequence 15, Appl
38	17	3.2	18	1	US-08-100-247-5	Sequence 5, Appl
39	17	3.2	18	1	US-08-483-146A-5	Sequence 5, Appl
40	17	3.2	18	1	US-08-232-513A-6	Sequence 6, Appl
41	17	3.2	18	1	US-08-484-594A-5	Sequence 5, Appl
42	17	3.2	18	3	US-09-231-159-20	Sequence 20, Appl
43	17	3.2	18	3	US-08-611-307-20	Sequence 20, Appl
44	17	3.2	18	4	US-09-148-030-2	Sequence 2, Appl
45	17	3.2	18	4	US-09-076-258A-5	Sequence 5, Appl

ALIGNMENTS

RESULT 1

US-09-352-548-1  
; Sequence 1, Application US/09352548  
; Patent No. 6500431  
; GENERAL INFORMATION:  
; APPLICANT: Gill, Parkash S.  
; APPLICANT: Parkash S. Gill, M.D., Inc.  
; TITLE OF INVENTION: No. 6500431el Inhibitors of Angiogenesis and Tumor Growth  
; FILE REFERENCE: 017986-000410US  
; CURRENT APPLICATION NUMBER: US/09/352,548  
; CURRENT FILING DATE: 1999-07-12  
; EARLIER APPLICATION NUMBER: US 60/092,647  
; EARLIER FILING DATE: 1998-07-13  
; NUMBER OF SEQ ID NOS: 59  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 1  
; LENGTH: 524  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
; FEATURE:  
; OTHER INFORMATION: proasoposin  
; NAME/KEY: PEPTIDE  
; LOCATION: (195)..(275)  
; OTHER INFORMATION: Saposin B  
; US-09-352-548-1

Query Match	100.0%;	Score 524;	DB 4;	Length 524;
Best Local Similarity	100.0%;	Pred. No. 0;		
Matches 524;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
Qy	1	MYALFLIASLGAALAGPVLGLKCTRGS	AVWCNVTASDCGAVKHCLQTVWNKPTVKS	60
Db	1	MYALFLIASLGAALAGPVLGLKCTRGS	AVWCNVTASDCGAVKHCLQTVWNKPTVKS	60
Qy	61	LPDICKDVVTAAGDMLKDNATTEELVYLEKTDWLPKPNMSASCKEIVDSYLPVILDI		120
Db	61	LPDICKDVVTAAGDMLKDNATTEELVYLEKTDWLPKPNMSASCKEIVDSYLPVILDI		120
Qy	121	IKGEMSRPGVCSALNLCESLQKHLAELNHQKLESNKIPELDMTEVVAPPMANIPLLLY		180
Db	121	IKGEMSRPGVCSALNLCESLQKHLAELNHQKLESNKIPELDMTEVVAPPMANIPLLLY		180
Qy	181	PDGPRSKPQKNDGVCDCIQMVTDTQTAVRNSTFFQALVEHVKCECDRLGFGMADI		240
Db	181	PDGPRSKPQKNDGVCDCIQMVTDTQTAVRNSTFFQALVEHVKCECDRLGFGMADI		240
Qy	241	CKNYISYSEIAIQMMHMQPKKEICALVGFCDEKVENPMQTLVPAKVASKNVIIPALELVE		300
Db	241	CKNYISYSEIAIQMMHMQPKKEICALVGFCDEKVENPMQTLVPAKVASKNVIIPALELVE		300

QY 301 PIKKEVPAKSDVYCEVCEFLVKEVTKLIDNNKTEKILDAFDKWCCKLPKSLSECEQV 360  
Db 301 PIKKEVPAKSDVYCEVCEFLVKEVTKLIDNNKTEKILDAFDKWCCKLPKSLSECEQV 360  
QY 361 VDTYGSILSILLEEVSPVLCVSMHLCSGTRLPALTTHVTQPKDGGFCEVCKKLVGYLD 420  
Db 361 VDTYGSILSILLEEVSPVLCVSMHLCSGTRLPALTTHVTQPKDGGFCEVCKKLVGYLD 420  
QY 421 RNLEKNSTKQEIILAALEKGSFLPDYPYQKQCDQFVAEYEPVLIEILVEVMDPSFVCLKIG 480  
Db 421 RNLEKNSTKQEIILAALEKGSFLPDYPYQKQCDQFVAEYEPVLIEILVEVMDPSFVCLKIG 480  
QY 481 ACPSAHKPLLGTEKCIWGPSYWCNTETAACQNAVEHCKRHWN 524  
Db 481 ACPSAHKPLLGTEKCIWGPSYWCNTETAACQNAVEHCKRHWN 524

## RESULT 2

US-08-100-247-2  
; Sequence 2, Application US/08100247  
; Patent No. 5571787  
; GENERAL INFORMATION:  
; APPLICANT: O'BRIEN, JOHN S.  
; APPLICANT: KISHIMOTO, YASUO  
; TITLE OF INVENTION: PROSAPOSIN AS A NEUROTROPHIC FACTOR  
; NUMBER OF SEQUENCES: 5  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: KNOBBE, MARTENS, OLSON AND BEAR  
; STREET: 620 NEWPORT CENTER DRIVE SIXTEENTH FLOOR  
; CITY: NEWPORT BEACH  
; STATE: CA  
; COUNTRY: USA  
; ZIP: 92660  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/100,247  
; FILING DATE: 19930730  
; CLASSIFICATION: 514  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Israelsen, Ned A.  
; REGISTRATION NUMBER: 29,655  
; REFERENCE/DOCKET NUMBER: OBRIEN.002A  
; TELEPHONE: 619-235-8550  
; TELEFAX: 619-235-0176  
; INFORMATION FOR SEQ ID NO: 2:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 523 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
; HYPOTHEICAL: NO  
; ANTI-SENSE: NO  
; FRAGMENT TYPE: N-terminal  
; IMMEDIATE SOURCE:  
; CLONE: PROSAPOSIN  
US-08-100-247-2

Query Match 42.7%; Score 224; DB 1; Length 523;  
Best Local Similarity 100.0%; Pred. No. 3.6e-207;  
Matches 224; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 301 PIKKEVPAKSDVYCEVCEFLVKEVTKLIDNNKTEKILDAFDKWCCKLPKSLSECEQV 360  
Db 300 PIKKEVPAKSDVYCEVCEFLVKEVTKLIDNNKTEKILDAFDKWCCKLPKSLSECEQV 359  
QY 361 VDTYGSILSILLEEVSPVLCVSMHLCSGTRLPALTTHVTQPKDGGFCEVCKKLVGYLD 420

Db 360 VDTYGSILSILLEEVSPVLCVSMHLCSGTRLPALTTHVTQPKDGGFCEVCKKLVGYLD 419  
QY 421 RNLEKNSTKQEIILAALEKGSFLPDYPYQKQCDQFVAEYEPVLIEILVEVMDPSFVCLKIG 480  
Db 420 RNLEKNSTKQEIILAALEKGSFLPDYPYQKQCDQFVAEYEPVLIEILVEVMDPSFVCLKIG 479  
QY 481 ACPSAHKPLLGTEKCIWGPSYWCNTETAACQNAVEHCKRHWN 524  
Db 480 ACPSAHKPLLGTEKCIWGPSYWCNTETAACQNAVEHCKRHWN 523

## RESULT 3

US-08-232-513A-3  
; Sequence 3, Application US/08232513A  
; Patent No. 5700909  
; GENERAL INFORMATION:  
; APPLICANT: O'Brien, John S.  
; TITLE OF INVENTION: Prosaposin and Cytokine-Derived Peptides  
; TITLE OF INVENTION: as Therapeutic Agents  
; NUMBER OF SEQUENCES: 20  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Campbell & Flores LLP  
; STREET: 4370 La Jolla Village Drive, Suite 700  
; CITY: San Diego  
; STATE: California  
; COUNTRY: USA  
; ZIP: 92122  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/232,513A  
; FILING DATE: 21-APR-1994  
; CLASSIFICATION: 514  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/100,247  
; FILING DATE: 30-JUL-1993  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Campbell, Cathryn A.  
; REGISTRATION NUMBER: 31,815  
; REFERENCE/DOCKET NUMBER: P-UD 1643  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (619) 535-9001  
; TELEFAX: (619) 535-8949  
; INFORMATION FOR SEQ ID NO: 3:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 523 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; FEATURE:  
; NAME/KEY: Protein  
; LOCATION: 1..523  
; OTHER INFORMATION: /label= Hum\_prosaposin  
US-08-232-513A-3

Query Match 42.7%; Score 224; DB 1; Length 523;  
Best Local Similarity 100.0%; Pred. No. 3.6e-207;  
Matches 224; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 301 PIKKEVPAKSDVYCEVCEFLVKEVTKLIDNNKTEKILDAFDKWCCKLPKSLSECEQV 360  
Db 300 PIKKEVPAKSDVYCEVCEFLVKEVTKLIDNNKTEKILDAFDKWCCKLPKSLSECEQV 359  
QY 361 VDTYGSILSILLEEVSPVLCVSMHLCSGTRLPALTTHVTQPKDGGFCEVCKKLVGYLD 420  
Db 360 VDTYGSILSILLEEVSPVLCVSMHLCSGTRLPALTTHVTQPKDGGFCEVCKKLVGYLD 419  
QY 421 RNLEKNSTKQEIILAALEKGSFLPDYPYQKQCDQFVAEYEPVLIEILVEVMDPSFVCLKIG 480

Db 420 RNLEKSTKOEILAALEKGCFLDPYQKQCDQFVAEYEPVLIEILVENVMDPSFVCLKIG 479

Qy 481 ACPSAHKPLLGTEKCIWGPSPYWCNTETAAQCNAVEHCKRHVWN 524

Db 480 ACPSAHKPLLGTEKCIWGPSPYWCNTETAAQCNAVEHCKRHVWN 523

RESULT 4

US-08-756-031-2

; Sequence 2, Application US/08756031

; Patent No. 6590074

; GENERAL INFORMATION:

; APPLICANT: O'BRIEN, JOHN S.

; APPLICANT: KISHIMOTO, YASUO

; TITLE OF INVENTION: PROSAPOSIN AS A NEUROTROPHIC FACTOR

; NUMBER OF SEQUENCES: 5

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: KNOBBE, MARTENS, OLSON AND BEAR

; STREET: 620 NEWPORT CENTER DRIVE SIXTEENTH FLOOR

; CITY: NEWPORT BEACH

; STATE: CA

; COUNTRY: USA

; ZIP: 92660

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: Patent In Release #1.0, Version #1.25

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/756,031

; FILING DATE: 26-NOV-1996

; CLASSIFICATION: 514

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: 08/100,247

; FILING DATE: 30-JUL-1993

; ATTORNEY/AGENT INFORMATION:

; NAME: Israelsen, Ned A.

; REGISTRATION NUMBER: 29,655

; REFERENCE/DOCKET NUMBER: O'BRIEN.002A

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: 619-235-8550

; TELEFAX: 619-235-0176

; INFORMATION FOR SEQ ID NO: 2:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 523 amino acids

; TYPE: amino acid

; STRANDEDNESS: single

; TOPOLOGY: linear

; MOLECULE TYPE: peptide

; HYPOTHETICAL: NO

; ANTI-SENSE: NO

; FRAGMENT TYPE: N-terminal

; IMMEDIATE SOURCE:

; CLONE: PROSAPOSIN

US-08-756-031-2

Query Match 42.7%; Score 224; DB 4; Length 523;

Best Local Similarity 100.0%; Pred. No. 3.6e-207;

Matches 224; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 301 PIKHEVPKSDVYCEVCEFLVKEVTKLIDNNKTEKEILDAFDKWCCKLPKSLSEECQEV 360

Db 300 PIKHEVPKSDVYCEVCEFLVKEVTKLIDNNKTEKEILDAFDKWCCKLPKSLSEECQEV 359

Qy 361 VDTGSSILSLILEVSPVCSMLHLCSTGTRLPALTTHVTPQKDGCFCEVCKLVGYLD 420

Db 360 VDTGSSILSLILEVSPVCSMLHLCSTGTRLPALTTHVTPQKDGCFCEVCKLVGYLD 419

Qy 421 RNLEKSTKOEILAALEKGCFLDPYQKQCDQFVAEYEPVLIEILVENVMDPSFVCLKIG 480

Db 420 RNLEKSTKOEILAALEKGCFLDPYQKQCDQFVAEYEPVLIEILVENVMDPSFVCLKIG 479

Qy 481 ACPSAHKPLLGTEKCIWGPSPYWCNTETAAQCNAVEHCKRHVWN 524

Db 480 ACPSAHKPLLGTEKCIWGPSPYWCNTETAAQCNAVEHCKRHVWN 523

RESULT 5

US-08-483-146A-2

; Sequence 2, Application US/08483146A

; Patent No. 5696080

; GENERAL INFORMATION:

; APPLICANT: O'Brien, John S.

; APPLICANT: Kishimoto, Yasuo

; TITLE OF INVENTION: PHARMACEUTICAL COMPOSITIONS

; TITLE OF INVENTION: COMPRISING PROSAPOSIN AND NEUROTROPHIC PEPTIDES DERIVED THEREFROM

; NUMBER OF SEQUENCES: 11

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Knobbe, Martens, Olson and Bear

; STREET: 620 Newport Center Blvd. 16th Floor

; CITY: Newport Beach

; STATE: CA

; COUNTRY: USA

; ZIP: 92660

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Diskette

; COMPUTER: IBM Compatible

; OPERATING SYSTEM: DOS

; SOFTWARE: FastSeq for Windows Version 2.0

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/483,146A

; FILING DATE: 07-JUN-1995

; CLASSIFICATION: 514

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER:

; FILING DATE:

; ATTORNEY/AGENT INFORMATION:

; NAME: Israelsen, Ned A.

; REGISTRATION NUMBER: 29,655

; REFERENCE/DOCKET NUMBER: MYELOS.002DV1

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: 619-235-8550

; TELEFAX: 619-235-0176

; TELEX:

; INFORMATION FOR SEQ ID NO: 2:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 523 amino acids

; TYPE: amino acid

; STRANDEDNESS: single

; TOPOLOGY: linear

; MOLECULE TYPE: peptide

; FRAGMENT TYPE: N-terminal

US-08-483-146A-2

Query Match 36.8%; Score 193; DB 1; Length 523;

Best Local Similarity 100.0%; Pred. No. 2.5e-177;

Matches 193; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MYALFLLASLLGALAGPVLGLKECTRGSAVCONVKTASDCGAVKHCLQTVWNKPTVKS 60

Db 1 MYALFLLASLLGALAGPVLGLKECTRGSAVCONVKTASDCGAVKHCLQTVWNKPTVKS 60

Qy 61 LPCDICKQVVVTAAGDMLKDNATBEEILVYLEKTCDWLPKPNMSASCKEIVDSYLPVILDI 120

Db 61 LPCDICKQVVVTAAGDMLKDNATBEEILVYLEKTCDWLPKPNMSASCKEIVDSYLPVILDI 120

Qy 121 IKGEMSRPGVCSALNLCESLQKHLELNHOKLESNKIPELDMTEVVAPFMANIPILLY 180

Db 121 IKGEMSRPGVCSALNLCESLQKHLELNHOKLESNKIPELDMTEVVAPFMANIPILLY 180

Qy 181 PQDGRSKPQPKD 193

Db 181 PQDGRSKPQPKD 193

## RESULT 6

US-08-484-594A-2  
; Sequence 2, Application US/08484594A  
; Patent No. 5714459  
; GENERAL INFORMATION:  
; APPLICANT: O'Brien, John S.  
; APPLICANT: Kishimoto, Yasuo  
; TITLE OF INVENTION: USE OF PROSAPIN AND NEUTROTROPIC PEPTIDES  
; TITLE OF INVENTION: DERIVED THEREFROM  
; NUMBER OF SEQUENCES: 11  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Knobbe, Martens, Olson and Bear  
; STREET: 620 Newport Center Drive, Sixteenth Floor  
; CITY: Newport Beach  
; STATE: CA  
; COUNTRY: USA  
; ZIP: 92660  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette  
; COMPUTER: IBM Compatible  
; OPERATING SYSTEM: DOS  
; SOFTWARE: FastSEQ for Windows Version 2.0  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/484,594A  
; FILING DATE: 07-JUN-1995  
; CLASSIFICATION: 514  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/100,247  
; FILING DATE: 30-JUL-1993  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Israelsen, Ned A.  
; REGISTRATION NUMBER: 29,655  
; REFERENCE/DOCKET NUMBER: MYELOS.002DV2  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 619-235-8550  
; TELEFAX: 619-235-0176  
; TELEX:  
; INFORMATION FOR SEQ ID NO: 2:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 523 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
; FRAGMENT TYPE: N-terminal  
US-08-484-594A-2

Query Match 36.8%; Score 193; DB 1; Length 523;  
Best Local Similarity 100.0%; Pred. No. 2.5e-177;  
Matches 193; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 1 MYALFLASLLGAALAGPVLGLKECTRGSAVWQNVKTASDCGAVKHCLQTVWNKPTVKS 60  
Db 1 MYALFLASLLGAALAGPVLGLKECTRGSAVWQNVKTASDCGAVKHCLQTVWNKPTVKS 60  
Qy 61 LPCDICKDVVTAAGDMLKDNATEEILVLEKTCDWLPKNMSASCKEIVDSYLPVLDI 120  
Db 61 LPCDICKDVVTAAGDMLKDNATEEILVLEKTCDWLPKNMSASCKEIVDSYLPVLDI 120  
Qy 121 IKGMSRPGVCSALNLCESLQKHLAELNHQKLESNKIPELDMTEVVAFFMANIPLLLY 180  
Db 121 IKGMSRPGVCSALNLCESLQKHLAELNHQKLESNKIPELDMTEVVAFFMANIPLLLY 180  
Qy 181 PDGPRSKPQPKD 193  
Db 181 PDGPRSKPQPKD 193

## RESULT 7

US-09-076-258A-2  
; Sequence 2, Application US/09076258A  
; Patent No. 6559124  
; GENERAL INFORMATION:

; APPLICANT: O'Brien, John S.  
; APPLICANT: Kishimoto, Yasuo  
; TITLE OF INVENTION: PHARMACEUTICAL COMPOSITIONS  
; TITLE OF INVENTION: COMPRISING PROSAPIN AND NEUTROTROPIC PEPTIDES DERIVED  
; TITLE OF INVENTION: THEREFROM  
; NUMBER OF SEQUENCES: 11  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Knobbe, Martens, Olson and Bear  
; STREET: 620 Newport Center Blvd. 16th Floor  
; CITY: Newport Beach  
; STATE: CA  
; COUNTRY: USA  
; ZIP: 92660  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette  
; COMPUTER: IBM Compatible  
; OPERATING SYSTEM: DOS  
; SOFTWARE: FastSEQ for Windows Version 2.0  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/076,258A  
; FILING DATE:  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/958,970  
; FILING DATE: 28-OCT-97  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/483,146  
; FILING DATE: 07-JUN-1995  
; APPLICATION NUMBER: 08/100,247  
; FILING DATE: 30-JUL-1993  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Israelsen, Ned A.  
; REGISTRATION NUMBER: 29,655  
; REFERENCE/DOCKET NUMBER: MYELOS.2DV1C2  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 619-235-8550  
; TELEFAX: 619-235-0176  
; TELEX:  
; INFORMATION FOR SEQ ID NO: 2:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 523 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
; FRAGMENT TYPE: N-terminal  
US-09-076-258A-2

Query Match 36.8%; Score 193; DB 4; Length 523;  
Best Local Similarity 100.0%; Pred. No. 2.5e-177;  
Matches 193; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 1 MYALFLASLLGAALAGPVLGLKECTRGSAVWQNVKTASDCGAVKHCLQTVWNKPTVKS 60  
Db 1 MYALFLASLLGAALAGPVLGLKECTRGSAVWQNVKTASDCGAVKHCLQTVWNKPTVKS 60  
Qy 61 LPCDICKDVVTAAGDMLKDNATEEILVLEKTCDWLPKNMSASCKEIVDSYLPVLDI 120  
Db 61 LPCDICKDVVTAAGDMLKDNATEEILVLEKTCDWLPKNMSASCKEIVDSYLPVLDI 120  
Qy 121 IKGMSRPGVCSALNLCESLQKHLAELNHQKLESNKIPELDMTEVVAFFMANIPLLLY 180  
Db 121 IKGMSRPGVCSALNLCESLQKHLAELNHQKLESNKIPELDMTEVVAFFMANIPLLLY 180  
Qy 181 PDGPRSKPQPKD 193  
Db 181 PDGPRSKPQPKD 193

## RESULT 8

US-09-352-548-2  
; Sequence 2, Application US/09352548  
; Patent No. 6500431

GENERAL INFORMATION:  
APPLICANT: Gill, Parkash S.  
APPLICANT: Parkash S. Gill, M.D., Inc.  
TITLE OF INVENTION: No. 6500431el Inhibitors of Angiogenesis and Tumor Growth  
FILE REFERENCE: 017986-000410US  
CURRENT APPLICATION NUMBER: US/09/352,548  
CURRENT FILING DATE: 1999-07-12  
EARLIER APPLICATION NUMBER: US 60/092,647  
EARLIER FILING DATE: 1998-07-13  
NUMBER OF SEQ ID NOS: 59  
SOFTWARE: PatentIn Ver. 2.1  
SEQ ID NO 2  
LENGTH: 81  
TYPE: PRT  
ORGANISM: Homo sapiens  
FEATURE:  
OTHER INFORMATION: Saposin B  
US-09-352-548-2

Query Match 15.5%; Score 81; DB 4; Length 81;  
Best Local Similarity 100.0%; Pred. No. 2.9e-70;  
Matches 81; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 195 GDVCDICQWTDIQTAVRTNSTFQALVEHVKECDRLGPGMADICKNYSIQYSEIAIQ 254  
Db 1 GDVCDICQWTDIQTAVRTNSTFQALVEHVKECDRLGPGMADICKNYSIQYSEIAIQ 60  
Qy 255 MMHMQPKEICALVGFCDVVK 275  
Db 61 MMHMQPKEICALVGFCDVVK 81

RESULT 9  
US-08-100-247-3  
Sequence 3, Application US/08100247  
Patent No. 5571787  
GENERAL INFORMATION:  
APPLICANT: O'BRIEN, JOHN S.  
APPLICANT: KISHIMOTO, YASUO  
TITLE OF INVENTION: PROSAPOSIN AS A NEUROTROPHIC FACTOR  
NUMBER OF SEQUENCES: 5  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: KNOBBE, MARTENS, OLSON AND BEAR  
STREET: 620 NEWPORT CENTER DRIVE SIXTEENTH FLOOR  
CITY: NEWPORT BEACH  
STATE: CA  
COUNTRY: USA  
ZIP: 92660  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/100,247  
FILING DATE: 19930730  
CLASSIFICATION: 514  
ATTORNEY/AGENT INFORMATION:  
NAME: Israel, Ned A.  
REGISTRATION NUMBER: 29,655  
REFERENCE/DOCKET NUMBER: O'BRIEN.002A  
TELEPHONE: 619-235-8550  
TELEFAX: 619-235-0176  
INFORMATION FOR SEQ ID NO: 3:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 80 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
HYPOTHETICAL: NO  
ANTI-SENSE: NO

FRAGMENT TYPE: internal  
IMMEDIATE SOURCE:  
CLONE: SAPOSIN C  
US-08-100-247-3  
Query Match 15.3%; Score 80; DB 1; Length 80;  
Best Local Similarity 100.0%; Pred. No. 2.6e-69;  
Matches 80; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 311 SDVYCEVCEFLVKEVTKLIDNNKTEKILDAFDMCKSLPKSLSEECQEVVDYTGSSILS 370  
Db 1 SDVYCEVCEFLVKEVTKLIDNNKTEKILDAFDMCKSLPKSLSEECQEVVDYTGSSILS 60  
Qy 371 ILLEEVSPELVCSMLHLCSG 390  
Db 61 ILLEEVSPELVCSMLHLCSG 80  
RESULT 10  
US-08-483-146A-3  
Sequence 3, Application US/08483146A  
Patent No. 5696080  
GENERAL INFORMATION:  
APPLICANT: O'Brien, John S.  
APPLICANT: Kishimoto, Yasuo  
TITLE OF INVENTION: PHARMACEUTICAL COMPOSITIONS  
TITLE OF INVENTION: COMPRISING PROSAPOSIN AND NEUROTROPHIC PEPTIDES DERIVED  
TITLE OF INVENTION: THEREFROM  
NUMBER OF SEQUENCES: 11  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Knobbe, Martens, Olson and Bear  
STREET: 620 Newport Center Blvd. 16th Floor  
CITY: Newport Beach  
STATE: CA  
COUNTRY: USA  
ZIP: 92660  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: DOS  
SOFTWARE: FastSeq for Windows Version 2.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/483,146A  
FILING DATE: 07-JUN-1995  
CLASSIFICATION: 514  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER:  
FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: Israel, Ned A.  
REGISTRATION NUMBER: 29,655  
REFERENCE/DOCKET NUMBER: MYELOS.002DV1  
TELEPHONE: 619-235-8550  
TELEFAX: 619-235-0176  
INFORMATION FOR SEQ ID NO: 3:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 80 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
FRAGMENT TYPE: internal  
US-08-483-146A-3  
Query Match 15.3%; Score 80; DB 1; Length 80;  
Best Local Similarity 100.0%; Pred. No. 2.6e-69;  
Matches 80; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 311 SDVYCEVCEFLVKEVTKLIDNNKTEKILDAFDMCKSLPKSLSEECQEVVDYTGSSILS 370  
Db 1 SDVYCEVCEFLVKEVTKLIDNNKTEKILDAFDMCKSLPKSLSEECQEVVDYTGSSILS 60

Qy 371 ILLEEVSPELVCSMLHLCSG 390  
Db 61 ILLEEVSPELVCSMLHLCSG 80

## RESULT 11

US-08-232-513A-4  
; Sequence 4, Application US/08232513A  
; Patent No. 5700909  
; GENERAL INFORMATION:  
; APPLICANT: O'Brien, John S.  
; TITLE OF INVENTION: Prosaposin and Cytokine-Derived Peptides  
; TITLE OF INVENTION: as Therapeutic Agents  
; NUMBER OF SEQUENCES: 20  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Campbell & Flores LLP  
; STREET: 4370 La Jolla Village Drive, Suite 700  
; CITY: San Diego  
; STATE: California  
; COUNTRY: USA  
; ZIP: 92122  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/232,513A  
; FILING DATE: 21-APR-1994  
; CLASSIFICATION: 514  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/100,247  
; FILING DATE: 30-JUL-1993  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Campbell, Cathryn A.  
; REGISTRATION NUMBER: 31,815  
; REFERENCE/DOCKET NUMBER: P-UD 1643  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (619) 535-9001  
; TELEFAX: (619) 535-8949  
; INFORMATION FOR SEQ ID NO: 4:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 80 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
; FEATURE:  
; NAME/KEY: Peptide  
; LOCATION: 1-80  
; OTHER INFORMATION: /label= Saposin\_C  
US-08-232-513A-4

Query Match 15.3%; Score 80; DB 1; Length 80;  
Best Local Similarity 100.0%; Pred. No. 2.6e-69;  
Matches 80; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 311 SDVYCEVCEFLVKEVTKLIDNNKTEKELIDAFDKMCKLPKSLSEECQEVVDYTGSSILS 370  
Db 1 SDVYCEVCEFLVKEVTKLIDNNKTEKELIDAFDKMCKLPKSLSEECQEVVDYTGSSILS 60

Qy 371 ILLEEVSPELVCSMLHLCSG 390  
Db 61 ILLEEVSPELVCSMLHLCSG 80

## RESULT 12

US-08-484-594A-3  
; Sequence 3, Application US/08484594A  
; Patent No. 5714459  
; GENERAL INFORMATION:  
; APPLICANT: O'Brien, John S.  
; APPLICANT: Kishimoto, Yasuo

; TITLE OF INVENTION: USE OF PROSAPOSIN AND NEUROTROPHIC PEPTIDES  
; TITLE OF INVENTION: DERIVED THEREFROM  
; NUMBER OF SEQUENCES: 11  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Knobbe, Martens, Olson and Bear  
; STREET: 620 Newport Center Drive, Sixteenth Floor  
; CITY: Newport Beach  
; STATE: CA  
; COUNTRY: USA  
; ZIP: 92660  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette  
; COMPUTER: IBM Compatible  
; OPERATING SYSTEM: DOS  
; SOFTWARE: FastSeq for Windows Version 2.0  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/484,594A  
; FILING DATE: 07-JUN-1995  
; CLASSIFICATION: 514  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/100,247  
; FILING DATE: 30-JUL-1993  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Israel, Ned A.  
; REGISTRATION NUMBER: 29,655  
; REFERENCE/DOCKET NUMBER: MYELOS.002DV2  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 619-235-8550  
; TELEFAX: 619-235-0176  
; TELEX:  
; INFORMATION FOR SEQ ID NO: 3:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 80 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; FRAGMENT TYPE: internal  
US-08-484-594A-3

Query Match 15.3%; Score 80; DB 1; Length 80;  
Best Local Similarity 100.0%; Pred. No. 2.6e-69;  
Matches 80; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 311 SDVYCEVCEFLVKEVTKLIDNNKTEKELIDAFDKMCKLPKSLSEECQEVVDYTGSSILS 370  
Db 1 SDVYCEVCEFLVKEVTKLIDNNKTEKELIDAFDKMCKLPKSLSEECQEVVDYTGSSILS 60

Qy 371 ILLEEVSPELVCSMLHLCSG 390  
Db 61 ILLEEVSPELVCSMLHLCSG 80

## RESULT 13

US-09-076-258A-3  
; Sequence 3, Application US/09076258A  
; Patent No. 6559124  
; GENERAL INFORMATION:  
; APPLICANT: O'Brien, John S.  
; APPLICANT: Kishimoto, Yasuo

; TITLE OF INVENTION: PHARMACEUTICAL COMPOSITIONS  
; TITLE OF INVENTION: COMPRISING PROSAPOSIN AND NEUROTROPHIC PEPTIDES DERIVED  
; NUMBER OF SEQUENCES: 11  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Knobbe, Martens, Olson and Bear  
; STREET: 620 Newport Center Blvd. 16th Floor  
; CITY: Newport Beach  
; STATE: CA  
; COUNTRY: USA  
; ZIP: 92660  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette



COMPUTER: IBM Compatible  
OPERATING SYSTEM: DOS  
SOFTWARE: FastSeq for Windows Version 2.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/076,258A  
FILING DATE:  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/958,970  
FILING DATE: 28-OCT-97  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/483,146  
FILING DATE: 07-JUN-1995  
APPLICATION NUMBER: 08/100,247  
FILING DATE: 30-JUL-1993  
ATTORNEY/AGENT INFORMATION:  
NAME: Israelsen, Ned A  
REGISTRATION NUMBER: 29,655  
REFERENCE/DOCKET NUMBER: MYELOS.2DVIC2  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 619-235-8550  
TELEFAX: 619-235-0176  
TELEX:  
INFORMATION FOR SEQ ID NO: 3:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 80 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
HYPOTHETICAL: NO  
ANTI-SENSE: NO  
FRAGMENT TYPE: internal  
IMMEDIATE SOURCE:  
CLONE: SAPOSIN C  
US-09-076-258A-3

Query Match 15.3%; Score 80; DB 4; Length 80;  
Best Local Similarity 100.0%; Pred. No. 2.6e-69;  
Matches 80; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 311 SDVYCEVCEFLVKEVTKLIDNNKTEKEILDADFDMCKSLPKSLSEECQEVVDITYGSSILS 370  
DB 1 SDVYCEVCEFLVKEVTKLIDNNKTEKEILDADFDMCKSLPKSLSEECQEVVDITYGSSILS 60  
QY 371 ILLEEVSPELVCSMLHLCSG 390  
DB 61 ILLEEVSPELVCSMLHLCSG 80

Query Match 15.3%; Score 80; DB 4; Length 80;  
Best Local Similarity 100.0%; Pred. No. 2.6e-69;  
Matches 80; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 311 SDVYCEVCEFLVKEVTKLIDNNKTEKEILDADFDMCKSLPKSLSEECQEVVDITYGSSILS 370  
DB 1 SDVYCEVCEFLVKEVTKLIDNNKTEKEILDADFDMCKSLPKSLSEECQEVVDITYGSSILS 60  
QY 371 ILLEEVSPELVCSMLHLCSG 390  
DB 61 ILLEEVSPELVCSMLHLCSG 80

RESULT 14  
US-08-756-031-3  
Sequence 3, Application US/08756031  
Patent No. 6590074  
GENERAL INFORMATION:  
APPLICANT: O'BRIEN, JOHN S.  
APPLICANT: KISHIMOTO, YASUO  
TITLE OF INVENTION: PROSAPOSIN AS A NEUROTROPHIC FACTOR  
NUMBER OF SEQUENCES: 5  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: KNOBBE, MARTENS, OLSON AND BEAR  
STREET: 620 NEWPORT CENTER DRIVE SIXTEENTH FLOOR  
CITY: NEWPORT BEACH  
STATE: CA  
COUNTRY: USA  
ZIP: 92660  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/756,031  
FILING DATE: 26-NOV-1996  
CLASSIFICATION: 514  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/100,247

FILING DATE: 30-JUL-1993  
ATTORNEY/AGENT INFORMATION:  
NAME: Israelsen, Ned A.  
REGISTRATION NUMBER: 29,655  
REFERENCE/DOCKET NUMBER: OBRIEN.002A  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 619-235-8550  
TELEFAX: 619-235-0176  
INFORMATION FOR SEQ ID NO: 3:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 80 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
HYPOTHETICAL: NO  
ANTI-SENSE: NO  
FRAGMENT TYPE: internal  
IMMEDIATE SOURCE:  
CLONE: SAPOSIN C  
US-08-756-031-3

Query Match 15.3%; Score 80; DB 4; Length 80;  
Best Local Similarity 100.0%; Pred. No. 2.6e-69;  
Matches 80; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 311 SDVYCEVCEFLVKEVTKLIDNNKTEKEILDADFDMCKSLPKSLSEECQEVVDITYGSSILS 370  
DB 1 SDVYCEVCEFLVKEVTKLIDNNKTEKEILDADFDMCKSLPKSLSEECQEVVDITYGSSILS 60  
QY 371 ILLEEVSPELVCSMLHLCSG 390  
DB 61 ILLEEVSPELVCSMLHLCSG 80

Query Match 15.3%; Score 80; DB 4; Length 80;  
Best Local Similarity 100.0%; Pred. No. 2.6e-69;  
Matches 80; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 311 SDVYCEVCEFLVKEVTKLIDNNKTEKEILDADFDMCKSLPKSLSEECQEVVDITYGSSILS 370  
DB 1 SDVYCEVCEFLVKEVTKLIDNNKTEKEILDADFDMCKSLPKSLSEECQEVVDITYGSSILS 60  
QY 371 ILLEEVSPELVCSMLHLCSG 390  
DB 61 ILLEEVSPELVCSMLHLCSG 80

RESULT 15  
US-08-584-671-15  
Sequence 15, Application US/08584671  
Patent No. 5910568  
GENERAL INFORMATION:  
APPLICANT: HAMMERSTEDT, ROY H, BARBATO, GUY F,  
APPLICANT: CRAMER, PALMER  
TITLE OF INVENTION: MOLECULE INVOLVED IN BINDING OF SPERM  
TITLE OF INVENTION: TO EGG SURFACES AND PROCEDURES FOR USE OF THIS MOLECULE  
TITLE OF INVENTION: TO ENHANCE OR DECREASE POTENTIAL FERTILITY  
NUMBER OF SEQUENCES: 16  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: INTELLECTUAL PROPERTY OFFICE, THE PENNSYLVANIA  
ADDRESSEE: STATE UNIVERSITY  
STREET: 113 TECHNOLOGY CENTER  
CITY: UNIVERSITY PARK  
STATE: PENNSYLVANIA  
COUNTRY: UNITED STATES OF AMERICA  
ZIP: 16802-7000  
COMPUTER READABLE FORM:  
MEDIUM TYPE: FLOPPY DISK  
COMPUTER: NEC 286  
OPERATING SYSTEM: DOS  
SOFTWARE: WORDPERFECT 5.1  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/584,671  
FILING DATE:  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: MONAHAN, THOMAS J  
REGISTRATION NUMBER: 29835  
REFERENCE/DOCKET NUMBER:  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 814-865-6277  
TELEFAX: 814-865-3591  
INFORMATION FOR SEQ ID NO: 15:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 80

;  
; TYPE: AMINO ACID  
; STRANDEDNESS: SINGLE  
; TOPOLOGY: UNKNOWN  
US-08-584-671-15

Query Match 14.3%; Score 75; DB 2; Length 80;  
Best Local Similarity 100.0%; Pred. No. 1.7e-64;  
Matches 75; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 134 ALNLCSLQKHIAELNHQKLESNKIPELDMTEVVVAPFMANIPLLLYPQDGPRSKPQPKD 193  
Db 6 ALNLCSLQKHIAELNHQKLESNKIPELDMTEVVVAPFMANIPLLLYPQDGPRSKPQPKD 65  
QY 194 NGDVCQDCIQWVTDI 208  
Db 66 NGDVCQDCIQWVTDI 80

Search completed: May 5, 2004, 13:35:19  
Job time : 16.595 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 5, 2004, 13:34:39 ; Search time 53.8081 Seconds  
(without alignments)  
2751.537 Million cell updates/sec

Title: US-09-743-684A-1

Perfect score: 524  
Sequence: 1 MYALFLASLLGALAGPVL.....NTETAACNAVEKRRHVN 524

Scoring table: OLIGO

Gapop 60.0 , Gapext 60.0  
Searched: 1586107 seqs, 282547505 residues

Word size : 0

Total number of hits satisfying chosen parameters: 1586107

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : A Geneseq\_29Jan04.\*  
1: Geneseqp1980s.\*  
2: Geneseqp1990s.\*  
3: Geneseqp2000s.\*  
4: Geneseqp2001s.\*  
5: Geneseqp2002s.\*  
6: Geneseqp2003as.\*  
7: Geneseqp2003bs.\*  
8: Geneseqp2004s.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	524	100.0	524	2 AAR70783	Aar70783 Prosaposin
2	524	100.0	524	2 AAW85652	Aaw85652 Human pro
3	524	100.0	524	3 AAY58716	Aay58716 Human pro
4	524	100.0	524	6 ABY79099	Abu79099 Lip-TAA b
5	524	100.0	524	6 ABY05200	Abu05200 Human exp
6	524	100.0	524	6 ABY05207	Abu05207 Human exp
7	524	100.0	524	6 ABY05203	Abu05203 Human exp
8	524	100.0	524	6 ABY07340	Abu07340 Human exp
9	524	100.0	524	6 ABY05216	Abu05216 Human exp
10	524	100.0	524	6 ABY05202	Abu05202 Human exp
11	524	100.0	524	6 ABY05214	Abu05214 Human exp
12	524	100.0	524	6 ABY05215	Abu05215 Human exp
13	524	100.0	524	6 ABY05199	Abu05199 Human exp
14	524	100.0	524	6 ABY05212	Abu05212 Human exp
15	524	100.0	524	6 ABY05213	Abu05213 Human exp
16	524	100.0	524	6 ABY05205	Abu05205 Human exp
17	423	80.7	524	6 ABY05208	Abu05208 Human exp
18	403	76.9	523	4 AAB31916	Aab31916 Amino aci
19	403	76.9	523	6 ABY05211	Abu05211 Human exp
20	284	54.2	527	4 AAB31915	Aab31915 Amino aci
21	284	54.2	527	5 ABP68602	Abp68602 Human pan
22	284	54.2	527	6 ABY79100	Abu79100 Lip-TAA b
23	284	54.2	527	6 ABY05204	Abu05204 Human exp
24	284	54.2	527	6 ABY05210	Abu05210 Human exp
25	265	50.6	385	6 ABR41750	Abu41750 Human DIT

ALIGNMENTS

RESULT 1

AAR70783  
ID AAR70783 standard; protein; 524 AA.

XX AC AAR70783;

XX DT 25-MAR-2003 (revised)

DT 30-AUG-1995 (first entry)

XX XX Prosaposin.

XX XX Saposin-C; neuron; myelination; nervous system; neuroblastoma;

KW neurotrophic peptide; multiple sclerosis; leukoencephalitis;

KW adrenal leukodystrophy; prosaposin.

XX XX Homo sapiens.

XX XX WO9503821-AI.

XX PD 09-FEB-1995.

XX PF 28-JUL-1994; 94WO-US008453.

XX PR 30-JUL-1993; 93US-00100247.

XX PR 21-APR-1994; 94US-00232513.

PA (OBRI/) O'BRIEN J S.

XX O'brien JS, Kishimoto Y;

PI WPI; 1995-082029/11.

XX N-PSDB; AAQ85355.

DR DR Stimulating neural cell out-growth and myelination - with pro-saposin,

PT saposin C or new neurotrophic peptide(s) from cytokine(s), for treating

PT nervous system diseases.

XX PS Disclosure; Page 30-32; 50pp; English.

XX CC The peptide given in AAR70773, corresponding to amino acids 8-29 of human

CC saposin-C (AAR70784), promotes neurite outgrowth in vitro. A consensus

CC sequence was determined by comparing the peptide with hematopoietic and

CC neurotrophic cytokines, and neurotrophic peptides (AAR70774-82) were

CC identified in the AB loop of human ciliary neurotrophic factor.

CC interleukins-6, -2, -3 and -gamma, erythropoietin and leukocyte

CC inhibitory factor, and in helix C of human interleukin-1-beta and

CC oncostatin-M. Prosaposin (AAR70783) and saposin-C also promoted nerve

CC cell myelination ex vivo. (Updated on 25-MAR-2003 to correct PN field.)

Abu05209 Human exp  
Abu05206 Human exp  
Aab39442 Human GEN  
Abu05201 Human exp  
Abg70166 Human pre  
Abu70504 Human adi  
Abu70799 Human adi  
Abu70422 Human adi  
Aar70784 Saposin-C  
Aaw85653 Human sap  
Aau05697 Human Sap  
Aau62252 Spingoli  
Aau05698 Human glu  
Aab31912 Amino aci  
Aab31929 Amino aci  
Aaw18584 Universal  
Aay33258 Human pro  
Aay62249 Spingoli  
Aau62250 Spingoli  
Aab67289 Human sap

CC (Updated on 25-MAR-2003 to correct PA field.) (Updated on 25-MAR-2003 to correct PI field.)

XX

SQ Sequence 524 AA;

Query Match 100.0%; Score 524; DB 2; Length 524;

Best Local Similarity 100.0%; Pred. No. 0;

Matches 524; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MYALFLLASLLGALAGPVLGLKECTRGSAVWQNVKTASDCGAVKHCLQTWNKPTVKS 60

DB 1 MYALFLLASLLGALAGPVLGLKECTRGSAVWQNVKTASDCGAVKHCLQTWNKPTVKS 60

QY 61 LPCDICKDVVTAAGDMLKDNATEEEILVYLEKTCMDLKPKNMSASCKEIVDSYLPVILDI 120

DB 61 LPCDICKDVVTAAGDMLKDNATEEEILVYLEKTCMDLKPKNMSASCKEIVDSYLPVILDI 120

QY 121 IKGEMSRPGEVCSALNLCESLQKHLAELNHQKLESNKIPELDMTEVVPFMANIPILLY 180

DB 121 IKGEMSRPGEVCSALNLCESLQKHLAELNHQKLESNKIPELDMTEVVPFMANIPILLY 180

QY 181 PDGGRSKPQPKDNGDVQCDCIQMTVDIQTAVRTNSTFVQALVEHVKEECDRLGPGWADI 240

DB 181 PDGGRSKPQPKDNGDVQCDCIQMTVDIQTAVRTNSTFVQALVEHVKEECDRLGPGWADI 240

QY 241 CKNYISQYSEIAIQMMHMQPKKEICALVGFCDVEKEMPQTLVPAKASKNVIIPALELVE 300

DB 241 CKNYISQYSEIAIQMMHMQPKKEICALVGFCDVEKEMPQTLVPAKASKNVIIPALELVE 300

QY 301 PIKKEVPAKSDVYCEVCEFLVKEVTKLIDNNKTEKEILDADFQKCKLPSLSEECQEV 360

DB 301 PIKKEVPAKSDVYCEVCEFLVKEVTKLIDNNKTEKEILDADFQKCKLPSLSEECQEV 360

QY 361 VDTYSSILSILLEVSPPELVCSMLHLCSTRLPALTVHTQPKDGGFCEVCKLVGYLD 420

DB 361 VDTYSSILSILLEVSPPELVCSMLHLCSTRLPALTVHTQPKDGGFCEVCKLVGYLD 420

QY 421 RNLEKNSFKQILAALEKGCFLDPYQKQCDQFVAEYEPVLBIELVMDPSFVCLKIG 480

DB 421 RNLEKNSFKQILAALEKGCFLDPYQKQCDQFVAEYEPVLBIELVMDPSFVCLKIG 480

QY 481 ACPSAHKPELLGTEKICWGPSWQNTETAACNAVEHCKRHVWN 524

DB 481 ACPSAHKPELLGTEKICWGPSWQNTETAACNAVEHCKRHVWN 524

RESULT 2

AAW85652

ID AAW85652 standard; protein; 524 AA.

XX

AC AAW85652;

XX

DT 19-JUL-1999 (first entry)

XX

DE Human prosaposin N-terminal peptide.

XX

KW Prosaposin; saposin; prosaptides; prosaposin receptor agonists; PRA;

KW peripheral nervous system; central nervous system; PNS; CNS; Akt; Bcl-2;

KW therapy; treatment; apoptosis; caspase; tumour necrosis factor; TNF;

KW cytokine; interferon gamma; IFN; inflammation; rheumatoid arthritis;

KW Crohn's disease; irritable bowel syndrome; asthma; cardiac infarction;

KW congestive heart failure; multiple sclerosis;

KW acute disseminated inflammatory leukoencephalitis;

KW progressive multifocal leukoencephalitis;

KW Parkinson's disease; amyotrophic lateral sclerosis; Huntington's disease;

KW ischemic heart disease; Guillain-Barre disease; alopecia; AIDS dementia;

KW cerebral malaria; HTLV; neuropathy;

KW inflammatory neurodegenerative disease; toxin-induced liver disease.

XX

OS Homo sapiens.

XX

OS

XX

PN W09912559-A1.

XX

PD 18-MAR-1999.

XX

PF 09-SEP-1998; 98WO-US019216.

XX

PR 09-SEP-1997; 97US-0058352P.

PR 04-JUN-1998; 98US-0088129P.

XX

PA (REGC) UNIV CALIFORNIA.

XX

PI O'Brien JS;

XX

DR WPI; 1999-229139/19.

XX

XX N-PSDB; AAX08488.

XX

XX Use of prosaposin receptor agonist.

PS Claim 7; Fig 2; 90pp; English.

XX

CC Prosaposin is a 70kDa glycoprotein which is proteolytically processed to generate saposins A, B, C and D, all of which are similar to each other and have a similar placement of six cysteines, a glycosylation site and conserved proline residues. Prosaposin, saposin C and prosaposin derived peptides (prosaptides), have therapeutic applications in promoting recovery after toxic, traumatic, myocardial ischaemic, degenerative and inherited lesions to the peripheral and central nervous system.

CC Prosaposin receptor agonists (PRAs) inhibit proinflammatory cytokine-induced apoptosis by activation of the Ser/Thr protein kinase Akt. Akt dissociates complexes of Bcl-2 family members, such as BAD-Bcl-2, releasing Bcl-2 and its family members which inhibit caspases, thereby inhibiting apoptosis. An additional mechanism whereby PRAs inhibit apoptosis is by blocking activation of JNK, a proapoptotic signaling component. Within several minutes after binding to the receptor, PRAs block JNK activation induced by tumor necrosis factor-alpha (TNF alpha). The activation of JNK by TNF alpha is another well known mechanism for TNF alpha-induced, as well as other proinflammatory cytokine-induced apoptosis. The method can be used for inhibiting apoptosis which is caspase-mediated or induced by a proinflammatory cytokine, for example TNF alpha or interferon-gamma. It can be used for inhibiting apoptosis associated with a disorder such as e.g. rheumatoid arthritis, Crohn's disease, irritable bowel syndrome, asthma, cardiac infarction, congestive heart failure, multiple sclerosis, acute disseminated inflammatory leukoencephalitis, progressive multifocal leukoencephalitis, Alzheimer's disease, Parkinson's disease, amyotrophic lateral sclerosis, Huntington's disease, ischemic heart disease, Guillain-Barre disease, traumatic brain injury, traumatic spinal cord injury, alopecia, AIDS dementia, cerebral malaria, HTLV, neuropathy, inflammatory neurodegenerative disease, and toxin-induced liver disease. This 524 N-terminal peptide of prosaposin also acts as a prosaposin receptor agonist.

XX

SQ Sequence 524 AA;

Query Match 100.0%; Score 524; DB 2; Length 524;

Best Local Similarity 100.0%; Pred. No. 0;

Matches 524; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MYALFLLASLLGALAGPVLGLKECTRGSAVWQNVKTASDCGAVKHCLQTWNKPTVKS 60

DB 1 MYALFLLASLLGALAGPVLGLKECTRGSAVWQNVKTASDCGAVKHCLQTWNKPTVKS 60

QY 61 LPCDICKDVVTAAGDMLKDNATEEEILVYLEKTCMDLKPKNMSASCKEIVDSYLPVILDI 120

DB 61 LPCDICKDVVTAAGDMLKDNATEEEILVYLEKTCMDLKPKNMSASCKEIVDSYLPVILDI 120

QY 121 IKGEMSRPGEVCSALNLCESLQKHLAELNHQKLESNKIPELDMTEVVPFMANIPILLY 180

DB 121 IKGEMSRPGEVCSALNLCESLQKHLAELNHQKLESNKIPELDMTEVVPFMANIPILLY 180

QY 181 PDGGRSKPQPKDNGDVQCDCIQMTVDIQTAVRTNSTFVQALVEHVKEECDRLGPGWADI 240

DB 181 PDGGRSKPQPKDNGDVQCDCIQMTVDIQTAVRTNSTFVQALVEHVKEECDRLGPGWADI 240

QY 241 CKNYISQYSEIAIQMMHMQPKKEICALVGFCDVEKEMPQTLVPAKASKNVIIPALELVE 300

XX

Db 241 CKNYISQYSEIAIQMMHMQPKSICALVGFCDVEKEMPMQTLVPAKVASKNVIPALELVE 300  
Qy 301 PIKKEVPAKSDVYCEVCEFLVKEVTKLIDNNKTEKELIDAFDKMCKSLPKSLSEECQEV 360  
Db 301 PIKKEVPAKSDVYCEVCEFLVKEVTKLIDNNKTEKELIDAFDKMCKSLPKSLSEECQEV 360  
Qy 361 VDTYGSILSILLEVSPELVCSMLHLCSTRLPALTVHVTQPKDGGFCVCKLVGYLD 420  
Db 361 VDTYGSILSILLEVSPELVCSMLHLCSTRLPALTVHVTQPKDGGFCVCKLVGYLD 420  
Qy 421 RNLEKNSTKQEIILAALEKGCFLPDYKQKODQFVAEYEPVLIEILVEVMDPSFVCLKIG 480  
Db 421 RNLEKNSTKQEIILAALEKGCFLPDYKQKODQFVAEYEPVLIEILVEVMDPSFVCLKIG 480  
Qy 481 ACPSAHKPLLGTEKCIWGPSYWCQNTETAACNAVEHCKRHVWN 524  
Db 481 ACPSAHKPLLGTEKCIWGPSYWCQNTETAACNAVEHCKRHVWN 524

RESULT 3  
AAV58716  
ID AAV58716 standard; protein; 524 AA.  
AC AAV58716;  
XX 25-APR-2000 (first entry)  
DT Human prosaposin.  
DE  
DE Prosapoin; saponin B; antiangiogenic; angiogenesis inhibitor;  
KW antitumour; antiproliferative; antimigratory; Kaposi's sarcoma; tumour;  
KW human; therapy.  
XX Homo sapiens.  
XX  
XX Key Location/Qualifiers  
FH Protein 195..:275  
FT /note= "mature saposin B"  
FT Peptide 195..:205  
FT /note= "specifically claimed antiangiogenic peptide of  
FT Claim 23"  
FT Peptide 196..:200  
FT /note= "specifically claimed antiangiogenic peptide of  
FT Claim 4"  
XX WO200002902-A1.  
XX 20-JAN-2000.  
XX 12-JUL-1999; 99WO-US015772.  
XX 13-JUL-1998; 98US-0092647P.  
XX (GILL/) GILL P S.  
XX Gill PS;  
XX WPI; 2000-171128/15.  
XX Saposin B derived peptides, useful as inhibitors of angiogenesis and  
PT tumor growth.  
XX  
XX Disclosure; Page 18; 78pp; English.  
XX The present sequence is that of human prosaposin, a precursor of saposin  
CC B. The invention is based on the discovery that saposin B, previously  
CC known to be involved in the hydrolysis of sphingolipids, has potent  
CC antiangiogenic and antitumour activity, and also has antiproliferative  
CC and antimigratory activity against endothelial cells. This activity is  
CC conserved in cryptic polypeptides as small as 5 amino acids (see AAV58684  
CC -715), which can be synthetically prepared and used in vitro or in vivo  
CC for the treatment of undesired angiogenesis and tumor growth, especially  
CC Kaposi's sarcoma (claimed). The polypeptides can also be used in

CC conjunction with cytotoxic moieties to selectively kill certain cell  
CC types, e.g. for treatment of cancer, angiofibroma, neovascular glaucoma,  
CC arteriovenous malformation, nonunion fracture, arthritis and other  
CC connective tissue disorders, Osler-Weber syndrome, atherosclerotic  
CC plaque, psoriasis, corneal graft neovascularization, pyogenic granuloma,  
CC retrolental fibroplasia, diabetic retinopathy, scleroderma, haemangioma,  
CC trachoma, vascular adhesions and hypertrophic scars  
XX  
SQ Sequence 524 AA;  
Query Match 100.0%; Score 524; DB 3; Length 524;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 524; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 1 MYALFLASLLGAALAGPVGLGKECTRGSAVWCQNVKTASDCGAVKHCLQTQVWKPVK 60  
Db 1 MYALFLASLLGAALAGPVGLGKECTRGSAVWCQNVKTASDCGAVKHCLQTQVWKPVK 60  
Qy 61 LPCDICKOVVTAAGDMLKDNATEEELVYLEKTCWMLPKPNMSASCKEIVDSYLPVLDI 120  
Db 61 LPCDICKOVVTAAGDMLKDNATEEELVYLEKTCWMLPKPNMSASCKEIVDSYLPVLDI 120  
Qy 121 IKGEMSRPGEVCSALNLCESLQKHLAELNHQKLESNKIPELDMTEVVAFFMANIPLLLY 180  
Db 121 IKGEMSRPGEVCSALNLCESLQKHLAELNHQKLESNKIPELDMTEVVAFFMANIPLLLY 180  
Qy 181 PODGPRSKPQKONGDVQCDCIQWVTDIQTAVRTNSTFFVQALVEHVKEECDRLGPGNADI 240  
Db 181 PODGPRSKPQKONGDVQCDCIQWVTDIQTAVRTNSTFFVQALVEHVKEECDRLGPGNADI 240  
Qy 241 CKNYISQYSEIAIQMMHMQPKSICALVGFCDVEKEMPMQTLVPAKVASKNVIPALELVE 300  
Db 241 CKNYISQYSEIAIQMMHMQPKSICALVGFCDVEKEMPMQTLVPAKVASKNVIPALELVE 300  
Qy 301 PIKKEVPAKSDVYCEVCEFLVKEVTKLIDNNKTEKELIDAFDKMCKSLPKSLSEECQEV 360  
Db 301 PIKKEVPAKSDVYCEVCEFLVKEVTKLIDNNKTEKELIDAFDKMCKSLPKSLSEECQEV 360  
Qy 361 VDTYGSILSILLEVSPELVCSMLHLCSTRLPALTVHVTQPKDGGFCVCKLVGYLD 420  
Db 361 VDTYGSILSILLEVSPELVCSMLHLCSTRLPALTVHVTQPKDGGFCVCKLVGYLD 420  
Qy 421 RNLEKNSTKQEIILAALEKGCFLPDYKQKODQFVAEYEPVLIEILVEVMDPSFVCLKIG 480  
Db 421 RNLEKNSTKQEIILAALEKGCFLPDYKQKODQFVAEYEPVLIEILVEVMDPSFVCLKIG 480  
Qy 481 ACPSAHKPLLGTEKCIWGPSYWCQNTETAACNAVEHCKRHVWN 524  
Db 481 ACPSAHKPLLGTEKCIWGPSYWCQNTETAACNAVEHCKRHVWN 524

RESULT 4  
ABU79099  
ID ABU79099 standard; protein; 524 AA.  
XX  
XX AC ABU79099;  
XX 18-JUN-2003 (first entry)  
DT  
DT Lip-TAA binding protein, Prosaposin.  
DE  
DE Superantigen; SAG; staphylococcal enterotoxin; tumour; cancer; apoptosis;  
KW gene therapy; mammalian cell receptor; cytostatic;  
KW tumour associated lipid; anergy; T cell; antigen presenting cell; APC;  
KW tumoricidal immunocyte; antitumour.  
XX Unidentified.  
XX OS US2002177551-A1.  
XX PN 28-NOV-2002.  
XX PD 30-MAY-2001; 2001US-00870759.  
XX PF

31-MAY-2000; 2000US-0208128P.  
(TERM/) TERMAN D S.  
Terman DS;  
WPI; 2003-361759/34.

A mammalian cell receptor, useful in the treatment of cancer by binding to tumor associated lipids where the binding induces energy or apoptosis in T cells and antigen presenting cells.

Disclosure; Page; 167pp; English.

The invention relates to a mammalian cell receptor, useful in the treatment of cancer, which binds to tumor associated lipids and induces energy or apoptosis in the T cells and antigen presenting cells (APCs). Also included are a mammalian cell useful in the treatment of cancer where the receptor which binds tumor associated lipids and induces cellular inactivation or death is deleted or functionally deactivated, producing (M1) a tumouricidal immunocyte population in vivo in a mammal (by allowing tumour associated lipids to contact immunocytes in which receptors for immunosuppressive fatty acids, ceramides, glycolipids, sphingolipids, glycosphingolipids, phosphosphingolipids, gangliosides, sialylated glycans, lipopeptides and proteoglycolipids are inactivated or deleted), a construct useful in the treatment of cancer comprising a superantigen (SAG) nucleotide inserted into a virus, a mammalian T cell useful in the treatment of cancer (where an adaptor protein which inhibits T cell activation by tumour associated antigens is deleted or functionally deactivated), a composition useful in the treatment of cancer (comprising a lipid raft conjugated to a superantigen), producing (M2) a tumouricidal immunocyte population ex vivo in a mammal (by allowing tumour associated lipids to contact immunocytes, in which receptors for the lipids are inactivated or deleted to produce a tumouricidal immunocyte population, and administering the tumouricidal APC to the host), producing (M3) a tumouricidal APC population ex vivo in a mammal (by allowing a tumour associated lipid to inactivate or deleted to produce a tumouricidally activated lipids are inactivated or deleted to produce a tumouricidally activated population, and administering APCs to the host), producing a tumouricidal T cell population ex vivo in a mammal (by allowing a tumour associated lipids to contact T cells, in which adaptor proteins, which inhibit T cell activation by tumour associated antigens, are deleted or functionally deactivated to produce a tumouricidal population of T cells, and administering the tumouricidally activated T cells to the host, or allowing a superantigen-lipid raft to contact T cells ex vivo, and administering the tumouricidally activated T cells to the host), treating (M5) cancer in a mammal (by administering a lipid binding molecule which binds immunosuppressive tumour associated lipids in vivo), producing (M6) a tumouricidal T cell population in vivo in a mammal (by allowing a tumour associated antigen to contact immunocytes in which adaptor proteins which inhibit T cell activation by tumour associated antigens are deleted or functionally deactivated) and producing (M7) a tumouricidal T cell population ex vivo in a mammal comprising allowing a superantigen-lipid raft conjugate to contact immunocytes in vivo. The receptors, methods and compositions are useful for treating cancers and tumours. Bacterial superantigens are useful for identifying a tumour sequence constructs with anti-tumour proteins or motifs. The present invention, which can be functionally deactivated in the method of the invention. Note: The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format from the US patent office website at "seqdata.uspto.gov/sequence.html?DocID=20020177551"

Sequence 524 AA;  
Query Match 100.0%; Score 524; DB 6; Length 524;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 524; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 MYALFLLASLGAALAGPVLGKTRGSVMQCNVKTASDCGAVKHCLQTVWVKPTVKS 60  
LPDICKDVVTAAGDMKONATEEELVLEKTCMDLPKNMSASCKEIVDSYLPVLDI 120  
LPDICKDVVTAAGDMKONATEEELVLEKTCMDLPKNMSASCKEIVDSYLPVLDI 120  
IKGEMSRPGVCSALNLCESLQKHAEHLNKHQLESNKIPELDMEVWAPPFMANIPLLY 180  
IKGEMSRPGVCSALNLCESLQKHAEHLNKHQLESNKIPELDMEVWAPPFMANIPLLY 180  
POGPRSKPQKNDGVCQDCIQMVTDTQTAVRTNSTFTVQALVEHVKEECDRGLPGMADI 240  
POGPRSKPQKNDGVCQDCIQMVTDTQTAVRTNSTFTVQALVEHVKEECDRGLPGMADI 240  
CKNYISQYSEIATQMMHMQPKKEICALVGFCDEVKEMPMQTLVPKAVSKNVIPALELYE 300  
CKNYISQYSEIATQMMHMQPKKEICALVGFCDEVKEMPMQTLVPKAVSKNVIPALELYE 300  
PIKKEVPAKSDVYCEVCFVKEVTKLIDNNKTEKILDAFDKMSCKLPKLSSECEQEV 360  
PIKKEVPAKSDVYCEVCFVKEVTKLIDNNKTEKILDAFDKMSCKLPKLSSECEQEV 360  
VDTYGSILSILLESVPELVCSMLHLCSTGLPALTVHVTQPKDGGFCVCKLVGYLD 420  
VDTYGSILSILLESVPELVCSMLHLCSTGLPALTVHVTQPKDGGFCVCKLVGYLD 420  
RNLEKNSTKQETLAALKEGCSFLPDYQKQDQFVAEVEPVLIEILVEWMDPSFVCLKIG 480  
RNLEKNSTKQETLAALKEGCSFLPDYQKQDQFVAEVEPVLIEILVEWMDPSFVCLKIG 480  
ACPSAHKPLLTGTEKINGSPSYWCNTETAACNAVEHCKRHWN 524  
ACPSAHKPLLTGTEKINGSPSYWCNTETAACNAVEHCKRHWN 524

RESULT 5  
ABU05200  
ID ABU05200 standard; protein; 524 AA.  
XX ABU05200;  
AC ABU05200;  
DT 29-JAN-2003 (first entry)  
XX Human expressed protein tag (EPT) #1866.  
XX  
XX Translational profiling; expressed protein tag; EPT; kinase; phosphatase;  
XX protease; protease inhibitor; transporter; cytoskeletal protein;  
XX receptor; transcription factor; cancer; MHC;  
XX major histocompatibility complex; myeloma; colon cancer; gastric cancer;  
XX adenocarcinoma; sarcoma; melanoma; lymphoma; leukaemia.  
XX Homo sapiens.  
XX  
XX WO200278524-A2.  
XX  
XX 10-OCT-2002.  
XX  
XX 28-MAR-2002; 2002WO-US009671.  
XX  
XX 28-MAR-2001; 2001US-0279495P.  
XX 21-MAY-2001; 2001US-0292544P.  
XX 08-AUG-2001; 2001US-0310801P.  
XX 01-OCT-2001; 2001US-0326370P.  
XX 04-DEC-2001; 2001US-0336780P.  
XX 20-FEB-2002; 2002US-0358985P.  
XX  
XX (ZYCO-) ZYCO INC.  
XX  
XX Chicx RM, Tomlinson AJ, Urban RG;  
XX  
XX WPI; 2003-040607/03.

PT New polypeptides (e.g. kinases, phosphatases, proteases, transporters,  
PT cytoskeletal proteins, receptors or transcription factors), useful for  
PT treating cancer, e.g. colon cancer, gastric cancer, sarcoma, lymphoma or  
PT leukemia.  
XX  
XX  
PS Example 2; SEQ ID NO 1866; 134pp; English.  
XX  
XX The invention describes a purified polypeptide, which comprises a  
CC fragment of a kinase, phosphatase, protease, protease inhibitor,  
CC transporter, cytoskeletal protein, receptor or transcription factor. The  
CC polypeptide is useful as an immunogenic composition for eliciting in a  
CC mammal an immunogenic response directed against any of the purified  
CC polypeptide. The purified polypeptide, or the antibody that binds to this  
CC polypeptide, is useful for treating cancer. The polypeptide is also  
CC useful for identifying compounds that binds to a naturally processed  
CC class I or class II MHC-binding polypeptide. The polypeptides and  
CC polynucleotides are particularly useful for treating or preventing  
CC myeloma, colon cancer, gastric cancer, adenocarcinoma, sarcoma, melanoma,  
CC lymphoma or leukaemia. These are also useful for screening agents for  
CC treating the above mentioned diseases. This sequence represents an  
CC expressed protein tag (EPT) isolated from human tissue for translational  
CC profiling. Note: This sequence does not appear in the printed  
CC specification but was obtained in electronic format directly from WIPO at  
CC ftp.wipo.int/pub/published\_pct\_sequences  
XX  
SQ Sequence 524 AA;  
Query Match 100.0%; Score 524; DB 6; Length 524;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 524; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 1 MYALLLASLILGALAGPVLGLKECTRGSAVWCVNKTASDCGAVKHCLQTVWNKPTVKS 60  
Db 1 MYALLLASLILGALAGPVLGLKECTRGSAVWCVNKTASDCGAVKHCLQTVWNKPTVKS 60  
Qy 61 LPCDICKDVVTAAGDMLKONATEEELVYLEKTCDWLPKPNMSASCKEIVDSYLPVILDI 120  
Db 61 LPCDICKDVVTAAGDMLKONATEEELVYLEKTCDWLPKPNMSASCKEIVDSYLPVILDI 120  
Qy 121 IKGMSRPGVCSALNLCESQKHLAELNHQKLESNKIPELDMTEVVAPPMANIPILLY 180  
Db 121 IKGMSRPGVCSALNLCESQKHLAELNHQKLESNKIPELDMTEVVAPPMANIPILLY 180  
Qy 181 PQDGRSKPQPKDNGDVQCQCIQMTDTQAVRTNSTFVQALVEHVKEECDRLGPGNADI 240  
Db 181 PQDGRSKPQPKDNGDVQCQCIQMTDTQAVRTNSTFVQALVEHVKEECDRLGPGNADI 240  
Qy 241 CKNYISQYSEIAIQMMHMQPKETCALVGFCDVEKEMPMQTLVPAKVASKNVIPALBELVE 300  
Db 241 CKNYISQYSEIAIQMMHMQPKETCALVGFCDVEKEMPMQTLVPAKVASKNVIPALBELVE 300  
Qy 301 PIKHEVPKSDVYCEVCEFLVKEVTKLIDNNKTEKELIDAFDKWCCKLPKSLSEECQEV 360  
Db 301 PIKHEVPKSDVYCEVCEFLVKEVTKLIDNNKTEKELIDAFDKWCCKLPKSLSEECQEV 360  
Qy 361 VDTYSSITSLILLEVSPPELVCSMLHLCSTGRLPALTVHTQKDGFCVCKKLGVGLD 420  
Db 361 VDTYSSITSLILLEVSPPELVCSMLHLCSTGRLPALTVHTQKDGFCVCKKLGVGLD 420  
Qy 421 RNLEKNSTKQIBILAALKEGCSFLPDPKQKCDQFVAEYEPVLIILVEVMDPSFVCLKIG 480  
Db 421 RNLEKNSTKQIBILAALKEGCSFLPDPKQKCDQFVAEYEPVLIILVEVMDPSFVCLKIG 480  
Qy 481 ACPSAHPLLGTEKICWGPSYWCQNTETAACNAVEHCKRHVN 524  
Db 481 ACPSAHPLLGTEKICWGPSYWCQNTETAACNAVEHCKRHVN 524  
RESULT 6  
ABU05207  
ID ABU05207 standard; protein; 524 AA.  
XX  
AC ABU05207;

XX 29-JAN-2003 (first entry)  
DT Human expressed protein tag (EPT) #1873.  
DE  
XX  
XX Translational profiling; expressed protein tag; EPT; kinase; phosphatase;  
KW protease; protease inhibitor; transporter; cytoskeletal protein;  
KW receptor; transcription factor; cancer; MHC;  
KW major histocompatibility complex; myeloma; colon cancer; gastric cancer;  
KW adenocarcinoma; sarcoma; melanoma; lymphoma; leukaemia.  
XX  
OS Homo sapiens.  
XX  
PN WO200278524-A2.  
XX  
PD 10-OCT-2002.  
XX  
PF 28-MAR-2002; 2002WO-US009671.  
XX  
PR 28-MAR-2001; 2001US-0279495P.  
XX 21-MAY-2001; 2001US-0292544P.  
PR 08-AUG-2001; 2001US-0310801P.  
PR 01-OCT-2001; 2001US-0326370P.  
PR 04-DEC-2001; 2001US-0336780P.  
PR 20-FEB-2002; 2002US-0358985P.  
XX  
XX (ZYCO-) ZYCOS INC.  
XX  
XX Chicx RM, Tomlinson AJ, Urban RG;  
XX  
XX WPI; 2003-040607/03.  
XX  
XX New polypeptides (e.g. kinases, phosphatases, proteases, transporters,  
PT cytoskeletal proteins, receptors or transcription factors), useful for  
PT treating cancer, e.g. colon cancer, gastric cancer, sarcoma, lymphoma or  
PT leukemia.  
XX  
XX Example 2; SEQ ID NO 1873; 134pp; English.  
XX  
XX The invention describes a purified polypeptide, which comprises a  
CC fragment of a kinase, phosphatase, protease, protease inhibitor,  
CC transporter, cytoskeletal protein, receptor or transcription factor. The  
CC polypeptide is useful as an immunogenic composition for eliciting in a  
CC mammal an immunogenic response directed against any of the purified  
CC polypeptide. The purified polypeptide, or the antibody that binds to this  
CC polypeptide, is useful for treating cancer. The polypeptide is also  
CC useful for identifying compounds that binds to a naturally processed  
CC class I or class II MHC-binding polypeptide. The polypeptides and  
CC polynucleotides are particularly useful for treating or preventing  
CC myeloma, colon cancer, gastric cancer, adenocarcinoma, sarcoma, melanoma,  
CC lymphoma or leukaemia. These are also useful for screening agents for  
CC treating the above mentioned diseases. This sequence represents an  
CC expressed protein tag (EPT) isolated from human tissue for translational  
CC profiling. Note: This sequence does not appear in the printed  
CC specification but was obtained in electronic format directly from WIPO at  
CC ftp.wipo.int/pub/published\_pct\_sequences  
XX  
SQ Sequence 524 AA;  
Query Match 100.0%; Score 524; DB 6; Length 524;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 524; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 1 MYALLLASLILGALAGPVLGLKECTRGSAVWCVNKTASDCGAVKHCLQTVWNKPTVKS 60  
Db 1 MYALLLASLILGALAGPVLGLKECTRGSAVWCVNKTASDCGAVKHCLQTVWNKPTVKS 60  
Qy 61 LPCDICKDVVTAAGDMLKONATEEELVYLEKTCDWLPKPNMSASCKEIVDSYLPVILDI 120  
Db 61 LPCDICKDVVTAAGDMLKONATEEELVYLEKTCDWLPKPNMSASCKEIVDSYLPVILDI 120  
Qy 121 IKGMSRPGVCSALNLCESQKHLAELNHQKLESNKIPELDMTEVVAPPMANIPILLY 180  
Db 121 IKGMSRPGVCSALNLCESQKHLAELNHQKLESNKIPELDMTEVVAPPMANIPILLY 180

Db 121 IKGEMSPGVCVCSALNLCESLQKHLAELNHQKLESNKIPELDMTEVVPFMANIPLLLY 180  
Qy 181 PDGPRSKPQKONGDVQCDCIQWVTDIQTAVRTNSTFVQALVEHVKECDRLGPGMADI 240  
Db 181 PDGPRSKPQKONGDVQCDCIQWVTDIQTAVRTNSTFVQALVEHVKECDRLGPGMADI 240  
Qy 241 CKNYISQYSIAIQMMHMQPKICALVGFCDSEKMPQTLVPKAVSKNVPALVELVE 300  
Db 241 CKNYISQYSIAIQMMHMQPKICALVGFCDSEKMPQTLVPKAVSKNVPALVELVE 300  
Qy 301 PIKKEVPAKSDVYCEVCEFLVKEVTKLIDNNKTEKILDAFDKWCCKLPSLSSECOEV 360  
Db 301 PIKKEVPAKSDVYCEVCEFLVKEVTKLIDNNKTEKILDAFDKWCCKLPSLSSECOEV 360  
Qy 361 VDTYGSLSILILEVSPVLCVSMHLCSGTRLPALTTHVTQPKDGGFCVCKLVGYLD 420  
Db 361 VDTYGSLSILILEVSPVLCVSMHLCSGTRLPALTTHVTQPKDGGFCVCKLVGYLD 420  
Qy 421 RNLEKNSTKQEIILAALEKGCSEFLPDYQKQCDQFVAEYEPVLIILVEVMDPSFVCLKIG 480  
Db 421 RNLEKNSTKQEIILAALEKGCSEFLPDYQKQCDQFVAEYEPVLIILVEVMDPSFVCLKIG 480  
Qy 481 ACPSAHKPLLGTEKICWGPSYWCNTTETAACNAVEHCKRHVWN 524  
Db 481 ACPSAHKPLLGTEKICWGPSYWCNTTETAACNAVEHCKRHVWN 524

RESULT 7  
ABU05203  
ID ABU05203 standard; protein; 524 AA.  
AC ABU05203;  
DT 29-JAN-2003 (first entry)  
XX Human expressed protein tag (EPT) #1869.  
DE Translational profiling; expressed protein tag; EPT; kinase; phosphatase;  
KW protease; protease inhibitor; transporter; cytoskeletal protein;  
KW receptor; transcription factor; cancer; MHC;  
KW major histocompatibility complex; myeloma; colon cancer; gastric cancer;  
KW adenocarcinoma; sarcoma; melanoma; lymphoma; leukaemia.  
OS Homo sapiens.  
XX  
XX WO200278524-A2.  
XX  
XX 10-OCT-2002.  
XX  
XX 28-MAR-2002; 2002WO-US009671.  
XX  
XX 28-MAR-2001; 2001US-0279495P.  
XX 21-MAY-2001; 2001US-0292544P.  
XX 08-AUG-2001; 2001US-0310801P.  
XX 01-OCT-2001; 2001US-0326370P.  
XX 04-DEC-2001; 2001US-0336780P.  
XX 20-FEB-2002; 2002US-0358985P.  
XX  
XX (ZYCO-) ZYCO INC.  
XX  
XX Chicx RM, Tomlinson AJ, Urban RG;  
XX  
XX WPI; 2003-040607/03.  
XX  
XX New polypeptides (e.g. kinases, phosphatases, proteases, transporters,  
XX cytoskeletal proteins, receptors or transcription factors), useful for  
XX treating cancer, e.g. colon cancer, gastric cancer, sarcoma, lymphoma or  
XX leukemia.  
XX  
XX Example 2; SEQ ID NO 1869; 134pp; English.  
XX  
XX The invention describes a purified polypeptide, which comprises a  
XX fragment of a kinase, phosphatase, protease, protease inhibitor,  
XX

CC transporter, cytoskeletal protein, receptor or transcription factor. The  
CC polypeptide is useful as an immunogenic composition for eliciting in a  
CC mammal an immunogenic response directed against any of the purified  
CC polypeptide. The purified polypeptide, or the antibody that binds to this  
CC polypeptide, is useful for treating cancer. The polypeptide is also  
CC useful for identifying compounds that binds to a naturally processed  
CC class I or class II MHC-binding polypeptide. The polypeptides and  
CC polynucleotides are particularly useful for treating or preventing  
CC myeloma, colon cancer, gastric cancer, adenocarcinoma, sarcoma, melanoma,  
CC lymphoma or leukaemia. These are also useful for screening agents for  
CC treating the above mentioned diseases. This sequence represents an  
CC expressed protein tag (EPT) isolated from human tissue for translational  
CC profiling. Note: This sequence does not appear in the printed  
CC specification but was obtained in electronic format directly from WIPO at  
CC ftp.wipo.int/pub/published\_pct\_sequences

XX Sequence 524 AA;

Query Match 100.0%; Score 524; DB 6; Length 524;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 524; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MYALFLLASLLGAALAGPVGLKECTRGSAVWQNVKTASDCGAVKHCLQTVWNKPTVKS 60  
Db 1 MYALFLLASLLGAALAGPVGLKECTRGSAVWQNVKTASDCGAVKHCLQTVWNKPTVKS 60  
Qy 61 LPCDICKVVTAAAGDMLKDNATEEBILVYLEKTDWLPKPNMSASCKEIVDSYLPVLDI 120  
Db 61 LPCDICKVVTAAAGDMLKDNATEEBILVYLEKTDWLPKPNMSASCKEIVDSYLPVLDI 120  
Qy 121 IKGEMSRPGVCVCSALNLCESLQKHLAELNHQKLESNKIPELDMTEVVPFMANIPLLLY 180  
Db 121 IKGEMSRPGVCVCSALNLCESLQKHLAELNHQKLESNKIPELDMTEVVPFMANIPLLLY 180  
Qy 181 PDGPRSKPQKONGDVQCDCIQWVTDIQTAVRTNSTFVQALVEHVKECDRLGPGMADI 240  
Db 181 PDGPRSKPQKONGDVQCDCIQWVTDIQTAVRTNSTFVQALVEHVKECDRLGPGMADI 240  
Qy 241 CKNYISQYSIAIQMMHMQPKICALVGFCDSEKMPQTLVPKAVSKNVPALVELVE 300  
Db 241 CKNYISQYSIAIQMMHMQPKICALVGFCDSEKMPQTLVPKAVSKNVPALVELVE 300  
Qy 301 PIKKEVPAKSDVYCEVCEFLVKEVTKLIDNNKTEKILDAFDKWCCKLPSLSSECOEV 360  
Db 301 PIKKEVPAKSDVYCEVCEFLVKEVTKLIDNNKTEKILDAFDKWCCKLPSLSSECOEV 360  
Qy 361 VDTYGSLSILILEVSPVLCVSMHLCSGTRLPALTTHVTQPKDGGFCVCKLVGYLD 420  
Db 361 VDTYGSLSILILEVSPVLCVSMHLCSGTRLPALTTHVTQPKDGGFCVCKLVGYLD 420  
Qy 421 RNLEKNSTKQEIILAALEKGCSEFLPDYQKQCDQFVAEYEPVLIILVEVMDPSFVCLKIG 480  
Db 421 RNLEKNSTKQEIILAALEKGCSEFLPDYQKQCDQFVAEYEPVLIILVEVMDPSFVCLKIG 480  
Qy 481 ACPSAHKPLLGTEKICWGPSYWCNTTETAACNAVEHCKRHVWN 524  
Db 481 ACPSAHKPLLGTEKICWGPSYWCNTTETAACNAVEHCKRHVWN 524

RESULT 8  
ABU07340  
ID ABU07340 standard; protein; 524 AA.  
XX  
XX AC ABU07340;  
XX  
XX DT 29-JAN-2003 (first entry)  
XX  
XX DE Human expressed protein tag (EPT) #2041.  
XX  
XX KW Translational profiling; expressed protein tag; EPT; kinase; phosphatase;  
KW protease; protease inhibitor; transporter; cytoskeletal protein;  
KW receptor; transcription factor; cancer; MHC;  
KW major histocompatibility complex; myeloma; colon cancer; gastric cancer;



```

KW adenocarcinoma; sarcoma; melanoma; lymphoma; leukaemia.
XX
OS Homo sapiens.
XX
PN WO200278524-A2.
XX
PD 10-OCT-2002.
XX
PF 28-MAR-2002; 2002WO-US009671.
XX
PR 28-MAR-2001; 2001US-0279495P.
PR 21-MAY-2001; 2001US-0292544P.
PR 08-AUG-2001; 2001US-0310801P.
PR 01-OCT-2001; 2001US-0326370P.
PR 04-DEC-2001; 2001US-0336780P.
PR 20-FEB-2002; 2002US-0358985P.
XX
PA (ZYCO-) ZYCOS INC.
XX
PI Chicx RM, Tomlinson AJ, Urban RG;
XX
DR WPI; 2003-040607/03.
XX
PT New polypeptides (e.g. kinases, phosphatases, proteases, transporters,
PT cytoskeletal proteins, receptors or transcription factors), useful for
PT treating cancer, e.g. colon cancer, gastric cancer, sarcoma, lymphoma or
PT leukaemia.
XX
PS Example 2; SEQ ID NO 2041; 134pp; English.
XX
CC The invention describes a purified polypeptide, which comprises a
CC fragment of a kinase, phosphatase, protease, protease inhibitor,
CC transporter, cytoskeletal protein, receptor or transcription factor. The
CC polypeptide is useful as an immunogenic composition for eliciting in a
CC mammal an immunogenic response directed against any of the purified
CC polypeptide. The purified polypeptide, or the antibody that binds to this
CC polypeptide, is useful for treating cancer. The polypeptide binds to this
CC useful for identifying compounds that binds to a naturally processed
CC class I or class II MHC-binding polypeptide. The polypeptides and
CC polynucleotides are particularly useful for treating or preventing
CC myeloma, colon cancer, gastric cancer, adenocarcinoma, sarcoma, melanoma,
CC lymphoma or leukaemia. These are also useful for screening agents for
CC treating the above mentioned diseases. This sequence represents an
CC expressed protein tag (EPT) isolated from human tissue for translational
CC profiling. Note: This sequence does not appear in the printed
CC specification but was obtained in electronic format directly from WIPD at
CC ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 524 AA;

Query Match 100.0%; Score 524; DB 6; Length 524;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 524; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MYALFLIASLLGALAGPVLGLKECTRGSAVWCNVTASDCGAVKHCLQTVWKNKPTVKS 60
DB 1 MYALFLIASLLGALAGPVLGLKECTRGSAVWCNVTASDCGAVKHCLQTVWKNKPTVKS 60
QY 61 LPDCICKDVVTAAGDMLKDNTATEEIIIVLEKTCMDLPKNMSASCKEIVDSYLPVILDI 120
DB 61 LPDCICKDVVTAAGDMLKDNTATEEIIIVLEKTCMDLPKNMSASCKEIVDSYLPVILDI 120
QY 121 IKGMSRPGVCSALNLCESLQKHLAELNHQKQLESNKIPELDMTEVVAPFMANIPILLY 180
DB 121 IKGMSRPGVCSALNLCESLQKHLAELNHQKQLESNKIPELDMTEVVAPFMANIPILLY 180
QY 181 PQDGPRSKPQKNGDVQCQCIQWVTDIQTAVRTNSTFVQALVEHVKKECDRLGPGNADI 240
DB 181 PQDGPRSKPQKNGDVQCQCIQWVTDIQTAVRTNSTFVQALVEHVKKECDRLGPGNADI 240
QY 241 CKNYISQYSEIATQMMHMPKEICALVGFCDVEKMPQTLVPKAVSKKNVIPALELVE 300
DB 241 CKNYISQYSEIATQMMHMPKEICALVGFCDVEKMPQTLVPKAVSKKNVIPALELVE 300

```

## RESULT 9

ABU05216

ID ABU05216 standard; protein; 524 AA.

XX AC ABU05216;

XX DT 29-JAN-2003 (first entry)

XX DE Human expressed protein tag (EPT) #1882.

XX KW Translational profiling; expressed protein tag; EPT; kinase; phosphatase;

XX KW protease; protease inhibitor; transporter; cytoskeletal protein;

XX KW receptor; transcription factor; cancer; MHC;

XX KW major histocompatibility complex; myeloma; colon cancer; gastric cancer;

XX KW adenocarcinoma; sarcoma; melanoma; lymphoma; leukaemia.

XX OS Homo sapiens.

XX PN WO200278524-A2.

XX PD 10-OCT-2002.

XX PF 28-MAR-2002; 2002WO-US009671.

XX PR 28-MAR-2001; 2001US-0279495P.

XX PR 21-MAY-2001; 2001US-0292544P.

XX PR 08-AUG-2001; 2001US-0310801P.

XX PR 01-OCT-2001; 2001US-0326370P.

XX PR 04-DEC-2001; 2001US-0336780P.

XX PR 20-FEB-2002; 2002US-0358985P.

XX PA (ZYCO-) ZYCOS INC.

XX PI Chicx RM, Tomlinson AJ, Urban RG;

XX DR WPI; 2003-040607/03.

XX PT New polypeptides (e.g. kinases, phosphatases, proteases, transporters,

XX PT cytoskeletal proteins, receptors or transcription factors), useful for

XX PT treating cancer, e.g. colon cancer, gastric cancer, sarcoma, lymphoma or

XX PT leukaemia.

XX PS Example 2; SEQ ID NO 1882; 134pp; English.

XX CC The invention describes a purified polypeptide, which comprises a

XX CC fragment of a kinase, phosphatase, protease, protease inhibitor,

XX CC transporter, cytoskeletal protein, receptor or transcription factor. The

XX CC polypeptide is useful as an immunogenic composition for eliciting in a

XX CC mammal an immunogenic response directed against any of the purified

XX CC polypeptide. The purified polypeptide, or the antibody that binds to this

XX CC polypeptide, is useful for treating cancer. The polypeptide is also

XX CC useful for identifying compounds that binds to a naturally processed

XX CC class I or class II MHC-binding polypeptide. The polypeptides and

XX CC polynucleotides are particularly useful for treating or preventing

XX CC myeloma, colon cancer, gastric cancer, adenocarcinoma, sarcoma, melanoma,

XX CC lymphoma, colon cancer, gastric cancer, adenocarcinoma, sarcoma, melanoma,

CC lymphoma or leukaemia. These are also useful for screening agents for  
CC treating the above mentioned diseases. This sequence represents an  
CC expressed protein tag (EPT) isolated from human tissue for translational  
CC profiling. Note: This sequence does not appear in the printed  
CC specification but was obtained in electronic format directly from WIPO at  
CC ftp.wipo.int/pub/published\_pct\_sequences  
XX  
SQ Sequence 524 AA;

```
Query Match      100.0%; Score 524; DB 6; Length 524;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 524; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MYALFLASLLGAALAGPVLGLKECTRGSAVWQNVKTASDCGAVKHCLQTVWVKPTVKS 60
Db 1 MYALFLASLLGAALAGPVLGLKECTRGSAVWQNVKTASDCGAVKHCLQTVWVKPTVKS 60
QY 61 LPDCICKDVVTAAGDMLKDNATEEEILVYLEKTCRWLKPKNMSASCKEIVDSYLPVLDI 120
Db 61 LPDCICKDVVTAAGDMLKDNATEEEILVYLEKTCRWLKPKNMSASCKEIVDSYLPVLDI 120
QY 121 IKGEMSRPGEVCSALNLCESLQKHLAELNHQKLESNKIPELDMTEVVPFMANIPLLY 180
Db 121 IKGEMSRPGEVCSALNLCESLQKHLAELNHQKLESNKIPELDMTEVVPFMANIPLLY 180
QY 181 PQDGRSKPQPKONGDVQDCIQMVTDTQTAVRTNSTFVQALVEHVKEECDRLGPGMADI 240
Db 181 PQDGRSKPQPKONGDVQDCIQMVTDTQTAVRTNSTFVQALVEHVKEECDRLGPGMADI 240
QY 241 CKNYISQYSEIAIQMMHMQPKIECALVGFCDVEKEMPQTLVPAKASKNVIIPALELVE 300
Db 241 CKNYISQYSEIAIQMMHMQPKIECALVGFCDVEKEMPQTLVPAKASKNVIIPALELVE 300
QY 301 PIKKEHVPKSDVYCEVCEFLVKEVTKLIDNNKTEKILDAFDKMSCKLPKSLSEECQEV 360
Db 301 PIKKEHVPKSDVYCEVCEFLVKEVTKLIDNNKTEKILDAFDKMSCKLPKSLSEECQEV 360
QY 361 VDTYGSLSILLEEVSPELVCSMLHCLSGTRLPALT VHVTQPKDGGFCVCKKLVGYLD 420
Db 361 VDTYGSLSILLEEVSPELVCSMLHCLSGTRLPALT VHVTQPKDGGFCVCKKLVGYLD 420
QY 421 RNLEKNSTKQIBLALEKSGSFLPDYQKQCDQFVAEYEPVLILVEVMDPSFVCLIKIG 480
Db 421 RNLEKNSTKQIBLALEKSGSFLPDYQKQCDQFVAEYEPVLILVEVMDPSFVCLIKIG 480
QY 481 ACPSAHKEPTEKICWGPSYWCQNTETAACNAVECHKRHWN 524
Db 481 ACPSAHKEPTEKICWGPSYWCQNTETAACNAVECHKRHWN 524
```

RESULT 10  
ABU05202  
ID ABU05202 standard; protein; 524 AA.  
XX AC ABU05202;  
XX XX  
XX 29-JAN-2003 (first entry)  
XX Human expressed protein tag (EPT) #1868.  
XX  
XX Translational profiling; expressed protein tag; EPT; kinase; phosphatase;  
XX protease; protease inhibitor; transporter; cytoskeletal protein;  
XX receptor; transcription factor; cancer; MHC;  
XX major histocompatibility complex; myeloma; colon cancer; gastric cancer;  
XX adenocarcinoma; sarcoma; melanoma; lymphoma; leukaemia.  
XX  
XX Homo sapiens.  
XX OS  
XX WO200278524-A2.  
XX PN  
XX 10-OCT-2002.  
XX PD  
XX 28-MAR-2002; 2002WO-US009671.  
XX PF

XX  
PR 28-MAR-2001; 2001US-0279495P.  
PR 21-MAY-2001; 2001US-0292544P.  
PR 08-AUG-2001; 2001US-0310801P.  
PR 01-OCT-2001; 2001US-0326370P.  
PR 04-DEC-2001; 2001US-0336780P.  
PR 20-FEB-2002; 2002US-0358985P.  
XX (ZYCO-) ZYCOS INC.  
XX Chicx RM, Tomlinson AJ, Urban RG;  
XX WPI; 2003-040607/03.  
XX  
XX New polypeptides (e.g. kinases, phosphatases, proteases, transporters,  
XX cytoskeletal proteins, receptors or transcription factors), useful for  
XX treating cancer, e.g. colon cancer, gastric cancer, sarcoma, lymphoma or  
XX leukemia.  
XX  
XX Example 2; SEQ ID NO 1868; 134pp; English.

XX The invention describes a purified polypeptide, which comprises a  
XX fragment of a kinase, phosphatase, protease, or transcription inhibitor,  
XX transporter, cytoskeletal protein, receptor or transcription factor. The  
XX polypeptide is useful as an immunogenic composition for eliciting in a  
XX mammal an immunogenic response directed against any of the purified  
XX polypeptide. The purified polypeptide, or the antibody that binds to this  
XX polypeptide, is useful for treating cancer. The polypeptide is also  
XX useful for identifying compounds that binds to a naturally processed  
XX class I or class II MHC-binding polypeptide. The polypeptides and  
XX polynucleotides are particularly useful for treating or preventing  
XX myeloma, colon cancer, gastric cancer, adenocarcinoma, sarcoma, melanoma,  
XX lymphoma or leukaemia. These are also useful for screening agents for  
XX treating the above mentioned diseases. This sequence represents an  
XX expressed protein tag (EPT) isolated from human tissue for translational  
XX profiling. Note: This sequence does not appear in the printed  
XX specification but was obtained in electronic format directly from WIPO at  
XX ftp.wipo.int/pub/published\_pct\_sequences  
XX  
XX Sequence 524 AA;

```
Query Match      100.0%; Score 524; DB 6; Length 524;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 524; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MYALFLASLLGAALAGPVLGLKECTRGSAVWQNVKTASDCGAVKHCLQTVWVKPTVKS 60
Db 1 MYALFLASLLGAALAGPVLGLKECTRGSAVWQNVKTASDCGAVKHCLQTVWVKPTVKS 60
QY 61 LPDCICKDVVTAAGDMLKDNATEEEILVYLEKTCRWLKPKNMSASCKEIVDSYLPVLDI 120
Db 61 LPDCICKDVVTAAGDMLKDNATEEEILVYLEKTCRWLKPKNMSASCKEIVDSYLPVLDI 120
QY 121 IKGEMSRPGEVCSALNLCESLQKHLAELNHQKLESNKIPELDMTEVVPFMANIPLLY 180
Db 121 IKGEMSRPGEVCSALNLCESLQKHLAELNHQKLESNKIPELDMTEVVPFMANIPLLY 180
QY 181 PQDGRSKPQPKONGDVQDCIQMVTDTQTAVRTNSTFVQALVEHVKEECDRLGPGMADI 240
Db 181 PQDGRSKPQPKONGDVQDCIQMVTDTQTAVRTNSTFVQALVEHVKEECDRLGPGMADI 240
QY 241 CKNYISQYSEIAIQMMHMQPKIECALVGFCDVEKEMPQTLVPAKASKNVIIPALELVE 300
Db 241 CKNYISQYSEIAIQMMHMQPKIECALVGFCDVEKEMPQTLVPAKASKNVIIPALELVE 300
QY 301 PIKKEHVPKSDVYCEVCEFLVKEVTKLIDNNKTEKILDAFDKMSCKLPKSLSEECQEV 360
Db 301 PIKKEHVPKSDVYCEVCEFLVKEVTKLIDNNKTEKILDAFDKMSCKLPKSLSEECQEV 360
QY 361 VDTYGSLSILLEEVSPELVCSMLHCLSGTRLPALT VHVTQPKDGGFCVCKKLVGYLD 420
Db 361 VDTYGSLSILLEEVSPELVCSMLHCLSGTRLPALT VHVTQPKDGGFCVCKKLVGYLD 420
```

Qy 421 RNLEKSTKQEIILAALEKGCFLPDPYKQCDQFVAEYEPVLIEILVEVMDPSFVCLKIG 480  
Db 421 RNLEKSTKQEIILAALEKGCFLPDPYKQCDQFVAEYEPVLIEILVEVMDPSFVCLKIG 480  
Qy 481 ACPSAHKPLLGTEKCIWGSPSYWQNTETAACQNAVEHCKRHVWN 524  
Db 481 ACPSAHKPLLGTEKCIWGSPSYWQNTETAACQNAVEHCKRHVWN 524

RESULT 11  
ABU05214  
ID ABU05214 standard; protein; 524 AA.  
XX AC ABU05214;  
XX DT 29-JAN-2003 (first entry)  
XX DE Human expressed protein tag (EPT) #1880.  
XX KW Translational profiling; expressed protein tag; EPT; kinase; phosphatase;  
KW protease; protease inhibitor; transporter; cytoskeletal protein;  
KW receptor; transcription factor; cancer; MHC;  
KW major histocompatibility complex; myeloma; colon cancer; gastric cancer;  
KW adenocarcinoma; sarcoma; melanoma; lymphoma; leukaemia.  
XX OS Homo sapiens.  
XX PN WO200278524-A2.  
XX PD 10-OCT-2002.  
XX PF 28-MAR-2002; 2002WO-US009671.  
XX PR 28-MAR-2001; 2001US-0279495P.  
PR 21-MAY-2001; 2001US-0292544P.  
PR 08-AUG-2001; 2001US-0310801P.  
PR 01-OCT-2001; 2001US-0326370P.  
PR 04-DEC-2001; 2001US-0336780P.  
PR 20-FEB-2002; 2002US-0358985P.  
XX PA (ZYCO-) ZYCOS INC.  
XX PI Chicx RM, Tomlinson AJ, Urban RG;  
XX DR WPI; 2003-040607/03.  
XX PT New polypeptides (e.g. kinases, phosphatases, proteases, transporters,  
PT cytoskeletal proteins, receptors or transcription factors), useful for  
PT treating cancer, e.g. colon cancer, gastric cancer, sarcoma, lymphoma or  
PT leukemia.  
XX PS Example 2; SEQ ID NO 1880; 134pp; English.  
XX CC The invention describes a purified polypeptide, which comprises a  
CC fragment of a kinase, phosphatase, protease, protease inhibitor,  
CC transporter, cytoskeletal protein, receptor or transcription factor. The  
CC polypeptide is useful as an immunogenic composition for eliciting in a  
CC mammal an immunogenic response directed against any of the purified  
CC polypeptide. The purified polypeptide, or the antibody that binds to this  
CC polypeptide, is useful for treating cancer. The polypeptide is also  
CC useful for identifying compounds that binds to a naturally processed  
CC class I or class II MHC-binding polypeptide. The polypeptides and  
CC polynucleotides are particularly useful for treating or preventing  
CC myeloma, colon cancer, gastric cancer, adenocarcinoma, sarcoma, melanoma,  
CC lymphoma or leukaemia. These are also useful for screening agents for  
CC treating the above mentioned diseases. This sequence represents an  
CC expressed protein tag (EPT) isolated from human tissue for translational  
CC profiling. Note: This sequence does not appear in the printed  
CC specification but was obtained in electronic format directly from WIPO at  
CC ftp.wipo.int/pub/published\_pct\_sequences  
XX Sequence 524 AA;

Query Match 100.0%; Score 524; DB 6; Length 524;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 524; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 1 MYALFLLSLLGAAALAGPVGLKECTRGSAVWQNVKTASDCGAVKHCLQTVNKPVTKS 60  
Db 1 MYALFLLSLLGAAALAGPVGLKECTRGSAVWQNVKTASDCGAVKHCLQTVNKPVTKS 60  
Qy 61 LPCDICDQVVTAGDMLKDNATEEEILVYLEKTCMDLKPKNMSASCKEIVDSYLPVLDI 120  
Db 61 LPCDICDQVVTAGDMLKDNATEEEILVYLEKTCMDLKPKNMSASCKEIVDSYLPVLDI 120  
Qy 121 IKGEMSRPGEVCSALNLCESLQKHLAELNHQKLESNKIPELDMTEVAPFMANIPLLLY 180  
Db 121 IKGEMSRPGEVCSALNLCESLQKHLAELNHQKLESNKIPELDMTEVAPFMANIPLLLY 180  
Qy 181 PDGPRSKPQKNDGVDVDCDQIOMVTDIOTAVRTNSTFVQALVEHVKEECDRLGPGWADI 240  
Db 181 PDGPRSKPQKNDGVDVDCDQIOMVTDIOTAVRTNSTFVQALVEHVKEECDRLGPGWADI 240  
Qy 241 CKNYISQYSEIAIQMMHMQPKKEICALVGFCDVEKEMPMQTLVPKAVSKNVIPALELVE 300  
Db 241 CKNYISQYSEIAIQMMHMQPKKEICALVGFCDVEKEMPMQTLVPKAVSKNVIPALELVE 300  
Qy 301 PIKKEVPKASDVYCEVCFVLKVEVTKLIDNNKTEKILDAFKMCSKLPKSLSEECQEV 360  
Db 301 PIKKEVPKASDVYCEVCFVLKVEVTKLIDNNKTEKILDAFKMCSKLPKSLSEECQEV 360  
Qy 361 VDTYSSIIISILLESVSPVLVCSMLHLCSTRLPALTVHVTQPKDGGFCVCKKLVGYLD 420  
Db 361 VDTYSSIIISILLESVSPVLVCSMLHLCSTRLPALTVHVTQPKDGGFCVCKKLVGYLD 420  
Qy 421 RNLEKSTKQEIILAALEKGCFLPDPYKQCDQFVAEYEPVLIEILVEVMDPSFVCLKIG 480  
Db 421 RNLEKSTKQEIILAALEKGCFLPDPYKQCDQFVAEYEPVLIEILVEVMDPSFVCLKIG 480  
Qy 481 ACPSAHKPLLGTEKCIWGSPSYWQNTETAACQNAVEHCKRHVWN 524  
Db 481 ACPSAHKPLLGTEKCIWGSPSYWQNTETAACQNAVEHCKRHVWN 524

RESULT 12  
ABU05215  
ID ABU05215 standard; protein; 524 AA.  
XX AC ABU05215;  
XX DT 29-JAN-2003 (first entry)  
XX DE Human expressed protein tag (EPT) #1881.  
XX KW Translational profiling; expressed protein tag; EPT; kinase; phosphatase;  
KW protease; protease inhibitor; transporter; cytoskeletal protein;  
KW receptor; transcription factor; cancer; MHC;  
KW major histocompatibility complex; myeloma; colon cancer; gastric cancer;  
KW adenocarcinoma; sarcoma; melanoma; lymphoma; leukaemia.  
XX OS Homo sapiens.  
XX PN WO200278524-A2.  
XX PD 10-OCT-2002.  
XX PF 28-MAR-2002; 2002WO-US009671.  
XX PR 28-MAR-2001; 2001US-0279495P.  
PR 21-MAY-2001; 2001US-0292544P.  
PR 08-AUG-2001; 2001US-0310801P.  
PR 01-OCT-2001; 2001US-0326370P.  
PR 04-DEC-2001; 2001US-0336780P.  
PR 20-FEB-2002; 2002US-0358985P.  
XX PA (ZYCO-) ZYCOS INC.

```
XX Chicz RM, Tomlinson AJ, Urban RG;
XX WPI; 2003-040607/03.
XX
XX New polypeptides (e.g. kinases, phosphatases, proteases, transporters,
XX cytoskeletal proteins, receptors or transcription factors), useful for
XX treating cancer, e.g. colon cancer, gastric cancer, sarcoma, lymphoma or
XX leukemia.
XX
XX Example 2; SEQ ID NO 1881; 134pp; English.
XX
XX The invention describes a purified polypeptide, which comprises a
XX fragment of a kinase, phosphatase, protease, protease inhibitor,
XX transporter, cytoskeletal protein, receptor or transcription factor. The
XX polypeptide is useful as an immunogenic composition for eliciting in a
XX mammal an immunogenic response directed against any of the purified
XX polypeptide. The purified polypeptide, or the antibody that binds to this
XX polypeptide, is useful for treating cancer. The polypeptide is also
XX useful for identifying compounds that binds to a naturally processed
XX class I or class II MHC-binding polypeptide. The polypeptides and
XX polynucleotides are particularly useful for treating or preventing
XX myeloma, colon cancer, gastric cancer, adenocarcinoma, sarcoma, melanoma,
XX lymphoma or leukaemia. These are also useful for screening agents for
XX treating the above mentioned diseases. This sequence represents an
XX expressed protein tag (EPT) isolated from human tissue for translational
XX profiling. Note: This sequence does not appear in the printed
XX specification but was obtained in electronic format directly from WIPO at
XX ftp.wipo.int/pub/published_pct_sequences
XX
XX Sequence 524 AA;
XX
XX Query Match 100.0%; Score 524; DB 6; Length 524;
XX Best Local Similarity 100.0%; Pred. No. 0;
XX Matches 524; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
QY 1 MYAFLLASLLGAALAGPVLGKCTRGSAVQCNVKTASDCGAVKHCLQTVWKNKPTVKS 60
Db 1 MYAFLLASLLGAALAGPVLGKCTRGSAVQCNVKTASDCGAVKHCLQTVWKNKPTVKS 60
QY 61 LPDCICKDVVTAAGDMLKDNATEEILLVLEKTCMDLPKPNMSASCKEIVDSYLPVILDI 120
Db 61 LPDCICKDVVTAAGDMLKDNATEEILLVLEKTCMDLPKPNMSASCKEIVDSYLPVILDI 120
QY 121 IKGMSRPGVCSALNLCESLQKHLAEINHQKLESNKIPELDMTEVVAFFMANIPILLY 180
Db 121 IKGMSRPGVCSALNLCESLQKHLAEINHQKLESNKIPELDMTEVVAFFMANIPILLY 180
QY 181 PQDGRSKPQPKNGDVQCDCIQMTVDIQTAVRTNSTFVQALVHVKEECDRILGPGMADI 240
Db 181 PQDGRSKPQPKNGDVQCDCIQMTVDIQTAVRTNSTFVQALVHVKEECDRILGPGMADI 240
QY 241 CKNVISQYSEIATIQMMHMQPKETCALVGFCDVEYKEMPMQTLVPAKVASKNVIPALELVE 300
Db 241 CKNVISQYSEIATIQMMHMQPKETCALVGFCDVEYKEMPMQTLVPAKVASKNVIPALELVE 300
QY 301 PIKHEVPKSDVYCEVFELVKEVTKLIDNNKTEKILDAFKMCSKLPKSLSESCQV 360
Db 301 PIKHEVPKSDVYCEVFELVKEVTKLIDNNKTEKILDAFKMCSKLPKSLSESCQV 360
QY 361 VDTYSSILSILLEEVSPELVCSMLHLCSTRLPALTVHTQPKDGGFCVCKLGYLD 420
Db 361 VDTYSSILSILLEEVSPELVCSMLHLCSTRLPALTVHTQPKDGGFCVCKLGYLD 420
QY 421 RNLEKNSTKGEILAALEKGSFLLPDYQKQCDQFVAEYEPVLIIEILVEMDFSVCLKIG 480
Db 421 RNLEKNSTKGEILAALEKGSFLLPDYQKQCDQFVAEYEPVLIIEILVEMDFSVCLKIG 480
QY 481 ACPSAHPLLGTEKCIWGPSYWCQNTETAAQCNAVEHCKRHVWN 524
Db 481 ACPSAHPLLGTEKCIWGPSYWCQNTETAAQCNAVEHCKRHVWN 524
```

```
RESULT 13
ABU05199
ID ABU05199 standard; protein; 524 AA.
XX
XX AC ABU05199;
XX
XX DT 29-JAN-2003 (first entry)
XX
XX DE Human expressed protein tag (EPT) #1865.
XX
XX KW Translational profiling; expressed protein tag; EPT; kinase; phosphatase;
XX protease; protease inhibitor; transporter; cytoskeletal protein;
XX receptor; transcription factor; cancer; MHC;
XX major histocompatibility complex; myeloma; colon cancer; gastric cancer;
XX adenocarcinoma; sarcoma; melanoma; lymphoma; leukaemia.
XX
XX OS Homo sapiens.
XX
XX PN W0200278524-A2.
XX
XX PD 10-OCT-2002.
XX
XX PF 28-MAR-2002; 2002WO-US009671.
XX
XX PR 28-MAR-2001; 2001US-0279495P.
XX
XX PR 21-MAY-2001; 2001US-0292544P.
XX
XX PR 08-AUG-2001; 2001US-0310801P.
XX
XX PR 01-OCT-2001; 2001US-0326370P.
XX
XX PR 04-DEC-2001; 2001US-0336780P.
XX
XX PR 20-FEB-2002; 2002US-0358985P.
XX
XX (ZYCO-) ZYCOS INC.
XX
XX Chicz RM, Tomlinson AJ, Urban RG;
XX
XX WPI; 2003-040607/03.
XX
XX New polypeptides (e.g. kinases, phosphatases, proteases, transporters,
XX cytoskeletal proteins, receptors or transcription factors), useful for
XX treating cancer, e.g. colon cancer, gastric cancer, sarcoma, lymphoma or
XX leukemia.
XX
XX Example 2; SEQ ID NO 1865; 134pp; English.
XX
XX The invention describes a purified polypeptide, which comprises a
XX fragment of a kinase, phosphatase, protease, protease inhibitor,
XX transporter, cytoskeletal protein, receptor or transcription factor. The
XX polypeptide is useful as an immunogenic composition for eliciting in a
XX mammal an immunogenic response directed against any of the purified
XX polypeptide. The purified polypeptide, or the antibody that binds to this
XX polypeptide, is useful for treating cancer. The polypeptide is also
XX useful for identifying compounds that binds to a naturally processed
XX class I or class II MHC-binding polypeptide. The polypeptides and
XX polynucleotides are particularly useful for treating or preventing
XX myeloma, colon cancer, gastric cancer, adenocarcinoma, sarcoma, melanoma,
XX lymphoma or leukaemia. These are also useful for screening agents for
XX treating the above mentioned diseases. This sequence represents an
XX expressed protein tag (EPT) isolated from human tissue for translational
XX profiling. Note: This sequence does not appear in the printed
XX specification but was obtained in electronic format directly from WIPO at
XX ftp.wipo.int/pub/published_pct_sequences
XX
XX Sequence 524 AA;
XX
XX Query Match 100.0%; Score 524; DB 6; Length 524;
XX Best Local Similarity 100.0%; Pred. No. 0;
XX Matches 524; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
QY 1 MYAFLLASLLGAALAGPVLGKCTRGSAVQCNVKTASDCGAVKHCLQTVWKNKPTVKS 60
Db 1 MYAFLLASLLGAALAGPVLGKCTRGSAVQCNVKTASDCGAVKHCLQTVWKNKPTVKS 60
QY 61 LPDCICKDVVTAAGDMLKDNATEEILLVLEKTCMDLPKPNMSASCKEIVDSYLPVILDI 120
```

Db 61 LPCDICKVVTAAAGDMLKDNATEEEILVYLEKTCQDMLPKFNMSASCKEIVDSYLPVILDI 120  
Qy 121 IKGEMSRPGEVCSALNLCESLQKHLAELNHQKLESNKIPELDMTEVVAFPMANIPILLY 180  
Db 121 IKGEMSRPGEVCSALNLCESLQKHLAELNHQKLESNKIPELDMTEVVAFPMANIPILLY 180  
Qy 181 PQDGRSKPQKONGDVQCQICQMTVDIQTAVRTNSTFFVQALVEHVKEECDRGLPGWADI 240  
Db 181 PQDGRSKPQKONGDVQCQICQMTVDIQTAVRTNSTFFVQALVEHVKEECDRGLPGWADI 240  
Qy 241 CKNYISQYSIAIQMMHMQPKEICALVGFCDVEKEMPMQTLVPAKVASKNVIPALELVE 300  
Db 241 CKNYISQYSIAIQMMHMQPKEICALVGFCDVEKEMPMQTLVPAKVASKNVIPALELVE 300  
Qy 301 PIKKEVPAKSDVYCEVCEFLVKEVTKLIDNNKTEKILDAFDMCKSKLPKSLSEECQEV 360  
Db 301 PIKKEVPAKSDVYCEVCEFLVKEVTKLIDNNKTEKILDAFDMCKSKLPKSLSEECQEV 360  
Qy 361 VDTYGSILSILLEEVSPELVCSMLHLCSTGTRLPALTVHVTQPKDGGFCVCKKLVGYLD 420  
Db 361 VDTYGSILSILLEEVSPELVCSMLHLCSTGTRLPALTVHVTQPKDGGFCVCKKLVGYLD 420  
Qy 421 RNLEKSTKQEIILAALEKSGSFLPDYQKQCDQFVAEYEPVLIIEILVEVMDPSFVCLKIG 480  
Db 421 RNLEKSTKQEIILAALEKSGSFLPDYQKQCDQFVAEYEPVLIIEILVEVMDPSFVCLKIG 480  
Qy 481 ACPSAHKPLLGTEKCIWGPSYWCQNTETAACQNAVEHCKRHVWN 524  
Db 481 ACPSAHKPLLGTEKCIWGPSYWCQNTETAACQNAVEHCKRHVWN 524

RESULT 14  
ABU05212  
ID ABU05212 standard; protein; 524 AA.  
AC ABU05212;  
XX  
DT 29-JAN-2003 (first entry)  
XX  
DE Human expressed protein tag (EPT) #1878.  
XX  
KW Translational profiling; expressed protein tag; EPT; kinase; phosphatase;  
KW protease; protease inhibitor; transporter; cytoskeletal protein;  
KW receptor; transcription factor; cancer; MHC;  
KW major histocompatibility complex; myeloma; colon cancer; gastric cancer;  
KW adenocarcinoma; sarcoma; melanoma; lymphoma; leukaemia.  
XX  
OS Homo sapiens.  
XX  
PN W0200278524-A2.  
XX  
PD 10-OCT-2002.  
XX  
PF 28-MAR-2002; 2002WO-US009671.  
XX  
PR 28-MAR-2001; 2001US-0279495P.  
PR 21-MAY-2001; 2001US-0292544P.  
PR 08-AUG-2001; 2001US-0310801P.  
PR 01-OCT-2001; 2001US-0326370P.  
PR 04-DEC-2001; 2001US-0336780P.  
PR 20-FEB-2002; 2002US-0358985P.  
XX  
PA (ZYCO-) ZYCO INC.  
XX  
PI Chicx RM, Tomlinson AJ, Urban RG;  
XX  
DR WPI; 2003-040607/03.  
XX  
PT New polypeptides (e.g. kinases, phosphatases, proteases, transporters,  
PT cytoskeletal proteins, receptors or transcription factors), useful for  
PT treating cancer, e.g. colon cancer, gastric cancer, sarcoma, lymphoma or  
PT leukemia.

XX  
PS  
XX  
CC The invention describes a purified polypeptide, which comprises a  
CC fragment of a kinase, phosphatase, protease, protease inhibitor,  
CC transporter, cytoskeletal protein, receptor or transcription factor. The  
CC polypeptide is useful as an immunogenic composition for eliciting in a  
CC mammal an immunogenic response directed against any of the purified  
CC polypeptide. The purified polypeptide, or the antibody that binds to this  
CC polypeptide, is useful for treating cancer. The polypeptide is also  
CC useful for identifying compounds that binds to a naturally processed  
CC class I or class II MHC-binding polypeptide. The polypeptides and  
CC polynucleotides are particularly useful for treating or preventing  
CC myeloma, colon cancer, gastric cancer, adenocarcinoma, sarcoma, melanoma,  
CC lymphoma or leukaemia. These are also useful for screening agents for  
CC treating the above mentioned diseases. This sequence represents an  
CC expressed protein tag (EPT) isolated from human tissue for translational  
CC profiling. Note: This sequence does not appear in the printed  
CC specification but was obtained in electronic format directly from WIPO at  
CC ftp.wipo.int/pub/published\_pct\_sequences  
XX  
SQ Sequence 524 AA;

Query Match 100.0%; Score 524; DB 6; Length 524;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 524; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 1 MYALFLLASLLGAALAGPVGLGKECTRGSAVWCNVKTASDCGAVKHCLQTVWNKPTVKS 60  
Db 1 MYALFLLASLLGAALAGPVGLGKECTRGSAVWCNVKTASDCGAVKHCLQTVWNKPTVKS 60  
Qy 61 LPCDICKVVTAAAGDMLKDNATEEEILVYLEKTCQDMLPKFNMSASCKEIVDSYLPVILDI 120  
Db 61 LPCDICKVVTAAAGDMLKDNATEEEILVYLEKTCQDMLPKFNMSASCKEIVDSYLPVILDI 120  
Qy 121 IKGEMSRPGEVCSALNLCESLQKHLAELNHQKLESNKIPELDMTEVVAFPMANIPILLY 180  
Db 121 IKGEMSRPGEVCSALNLCESLQKHLAELNHQKLESNKIPELDMTEVVAFPMANIPILLY 180  
Qy 181 PQDGRSKPQKONGDVQCQICQMTVDIQTAVRTNSTFFVQALVEHVKEECDRGLPGWADI 240  
Db 181 PQDGRSKPQKONGDVQCQICQMTVDIQTAVRTNSTFFVQALVEHVKEECDRGLPGWADI 240  
Qy 241 CKNYISQYSIAIQMMHMQPKEICALVGFCDVEKEMPMQTLVPAKVASKNVIPALELVE 300  
Db 241 CKNYISQYSIAIQMMHMQPKEICALVGFCDVEKEMPMQTLVPAKVASKNVIPALELVE 300  
Qy 301 PIKKEVPAKSDVYCEVCEFLVKEVTKLIDNNKTEKILDAFDMCKSKLPKSLSEECQEV 360  
Db 301 PIKKEVPAKSDVYCEVCEFLVKEVTKLIDNNKTEKILDAFDMCKSKLPKSLSEECQEV 360  
Qy 361 VDTYGSILSILLEEVSPELVCSMLHLCSTGTRLPALTVHVTQPKDGGFCVCKKLVGYLD 420  
Db 361 VDTYGSILSILLEEVSPELVCSMLHLCSTGTRLPALTVHVTQPKDGGFCVCKKLVGYLD 420  
Qy 421 RNLEKSTKQEIILAALEKSGSFLPDYQKQCDQFVAEYEPVLIIEILVEVMDPSFVCLKIG 480  
Db 421 RNLEKSTKQEIILAALEKSGSFLPDYQKQCDQFVAEYEPVLIIEILVEVMDPSFVCLKIG 480  
Qy 481 ACPSAHKPLLGTEKCIWGPSYWCQNTETAACQNAVEHCKRHVWN 524  
Db 481 ACPSAHKPLLGTEKCIWGPSYWCQNTETAACQNAVEHCKRHVWN 524

RESULT 15  
ABU05213  
ID ABU05213 standard; protein; 524 AA.  
XX  
AC ABU05213;  
XX  
DT 29-JAN-2003 (first entry)  
XX  
DE Human expressed protein tag (EPT) #1879.

XX Translational profiling; expressed protein tag; EPT; kinase; phosphatase;  
 KW protease; protease inhibitor; transporter; cytoskeletal protein;  
 KW receptor; transcription factor; cancer; MHC;  
 KW major histocompatibility complex; myeloma; colon cancer; gastric cancer;  
 KW adenocarcinoma; sarcoma; melanoma; lymphoma; leukaemia.  
 XX Homo sapiens.  
 XX OS  
 XX WC200278524-A2.  
 XX PD  
 XX 10-OCT-2002.  
 XX PF  
 XX 28-MAR-2002; 2002WO-US009671.  
 XX 28-MAR-2001; 2001US-0279495P.  
 PR 21-MAY-2001; 2001US-0292544P.  
 PR 08-AUG-2001; 2001US-0310801P.  
 PR 01-OCT-2001; 2001US-0326370P.  
 PR 04-DEC-2001; 2001US-0336780P.  
 PR 20-FEB-2002; 2002US-0358985P.  
 XX (ZYCO-) ZYCOS INC.  
 XX Chicx RM, Tomlinson AJ, Urban RG;  
 XX WPI; 2003-040607/03.  
 XX New polypeptides (e.g. kinases, phosphatases, proteases, transporters,  
 PT cytoskeletal proteins, receptors or transcription factors), useful for  
 PT treating cancer, e.g. colon cancer, gastric cancer, sarcoma, lymphoma or  
 PT leukemia.  
 XX Example 2; SEQ ID NO 1879; 134pp; English.  
 XX The invention describes a purified polypeptide, which comprises a  
 CC fragment of a kinase, phosphatase, protease, protease inhibitor,  
 CC transporter, cytoskeletal protein, receptor or transcription factor. The  
 CC polypeptide is useful as an immunogenic composition for eliciting in a  
 CC mammal an immunogenic response directed against any of the purified  
 CC polypeptide. The purified polypeptide, or the antibody that binds to this  
 CC polypeptide, is useful for treating cancer. The polypeptide is also  
 CC useful for identifying compounds that binds to a naturally processed  
 CC class I or class II MHC-binding polypeptide. The polypeptides and  
 CC polynucleotides are particularly useful for treating or preventing  
 CC myeloma, colon cancer, gastric cancer, adenocarcinoma, sarcoma, melanoma,  
 CC lymphoma or leukaemia. These are also useful for screening agents for  
 CC treating the above mentioned diseases. This sequence represents an  
 CC expressed protein tag (EPT) isolated from human tissue for translational  
 CC profiling. Note: This sequence does not appear in the printed  
 CC specification but was obtained in electronic format directly from WIPO at  
 CC ftp.wipo.int/pub/published\_pct\_sequences  
 XX Sequence 524 AA;  
 SQ

Query Match  
 Best Local Similarity 100.0%; Score 524; DB 6; Length 524;  
 Matches 524; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MYALFLASLLGALAGPVLGLKECTRGSAVWCQNVKTASDCGAVKHCLQTVWNKPTVKS 60  
 DB 1 MYALFLASLLGALAGPVLGLKECTRGSAVWCQNVKTASDCGAVKHCLQTVWNKPTVKS 60  
 QY 61 LPDCIDKDVVTAAGDMLKNATEEEILVLEKTCDWLPKPNMSASCKEIVDSYLPVILDI 120  
 DB 61 LPDCIDKDVVTAAGDMLKNATEEEILVLEKTCDWLPKPNMSASCKEIVDSYLPVILDI 120  
 QY 121 IKGEMSRPGVCSALNLCESLQKHAEHLNKHQKLESNKIPELDMTEVAVFNNIPLILY 180  
 DB 121 IKGEMSRPGVCSALNLCESLQKHAEHLNKHQKLESNKIPELDMTEVAVFNNIPLILY 180  
 QY 181 PQDGPRSKPOPKONGDVCDICQWVTDIQTAVRTNSTFVQALVEHVKECDRLGPGMADI 240  
 DB 181 PQDGPRSKPOPKONGDVCDICQWVTDIQTAVRTNSTFVQALVEHVKECDRLGPGMADI 240

Db 181 PQDGPRSKPOPKONGDVCDICQWVTDIQTAVRTNSTFVQALVEHVKECDRLGPGMADI 240  
 QY 241 CKNYISOYSIAIQMMHMQPKKEICALVGFCDVEKMPQTLVPAKVASKNVIPALELVE 300  
 Db 241 CKNYISOYSIAIQMMHMQPKKEICALVGFCDVEKMPQTLVPAKVASKNVIPALELVE 300  
 QY 301 PIKHEVPKSDVYCEVCEFLVKEVTKLIDNNKTEKEILDAFDMGSKLPKSLSEECQEV 360  
 Db 301 PIKHEVPKSDVYCEVCEFLVKEVTKLIDNNKTEKEILDAFDMGSKLPKSLSEECQEV 360  
 QY 361 VDTYGSILSILLEVSPELVCSMLHLCSTRLPALTVHVTOPKDGGFCEVCKLVGYLD 420  
 Db 361 VDTYGSILSILLEVSPELVCSMLHLCSTRLPALTVHVTOPKDGGFCEVCKLVGYLD 420  
 QY 421 RNLEKNSTKQEIILAALEKGCSEFLPDYQKQCDQFVAEYEPVLIIEILVEVMDPSFVCLKIG 480  
 Db 421 RNLEKNSTKQEIILAALEKGCSEFLPDYQKQCDQFVAEYEPVLIIEILVEVMDPSFVCLKIG 480  
 QY 481 ACPSAHKPLLLGTEKCTIWGSPSYWCQNTETAACQNAVEHCHKRVWN 524  
 Db 481 ACPSAHKPLLLGTEKCTIWGSPSYWCQNTETAACQNAVEHCHKRVWN 524

Search completed: May 5, 2004, 13:45:40  
 Job time : 54.8081 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 5, 2004, 13:34:39 ; Search time 53.1919 Seconds  
(without alignments)  
2751.537 Million cell updates/sec

Title: US-09-743-684A-1\_COPY\_7\_524

Perfect score: 518

Sequence: 1 LASLLGALAGPVLGLKECT.....NTETAAQCNAVEHCKRHVWN 518

Scoring table: OLIGO  
Gapop 60.0 , Gapext 60.0

Searched: 1586107 seqs, 282547505 residues

Word size : 0

Total number of hits satisfying chosen parameters: 1586107

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : A\_Geneseq\_29Jan04.\*

- 1: geneseqp1980s.\*
- 2: geneseqp1990s.\*
- 3: geneseqp2000s.\*
- 4: geneseqp2001s.\*
- 5: geneseqp2002s.\*
- 6: geneseqp2003as.\*
- 7: geneseqp2003bs.\*
- 8: geneseqp2004s.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	518	100.0	524	2	AAR70783 Prosaposin
2	518	100.0	524	2	AAW85652 Human pro
3	518	100.0	524	3	AAV58716 Human pro
4	518	100.0	524	6	ABU79099 Lip-TAA b
5	518	100.0	524	6	ABU05200 Human exp
6	518	100.0	524	6	ABU05207 Human exp
7	518	100.0	524	6	ABU05203 Human exp
8	518	100.0	524	6	ABU07340 Human exp
9	518	100.0	524	6	ABU05216 Human exp
10	518	100.0	524	6	ABU05202 Human exp
11	518	100.0	524	6	ABU05214 Human exp
12	518	100.0	524	6	ABU05215 Human exp
13	518	100.0	524	6	ABU05199 Human exp
14	518	100.0	524	6	ABU05212 Human exp
15	518	100.0	524	6	ABU05213 Human exp
16	518	100.0	524	6	ABU05205 Human exp
17	417	80.5	524	6	ABU05208 Human exp
18	397	76.6	523	4	AAH31916 Amino aci
19	397	76.6	523	6	ABU05211 Human exp
20	278	53.7	527	4	AAH31915 Amino aci
21	278	53.7	527	5	ABP68602 Human pan
22	278	53.7	527	6	ABU79100 Lip-TAA b
23	278	53.7	527	6	ABU05204 Human exp
24	278	53.7	527	6	ABU05210 Human exp
25	265	51.2	385	6	ABR41750 Human DIT

26	265	51.2	526	6	ABU05209 Human exp
27	265	51.2	527	6	ABU05206 Human exp
28	254	49.0	479	6	ABR39442 Human GEN
29	210	40.5	210	6	ABU05201 Human exp
30	209	40.3	209	5	ABG70166 Human pre
31	153	29.5	153	6	ABU70504 Human adi
32	153	29.5	153	6	ABU70799 Human adi
33	129	24.9	129	6	ABU70422 Human adi
34	80	15.4	80	2	AAR70784 Saposin-C
35	80	15.4	80	2	AAW85653 Human sap
36	80	15.4	80	4	AAU05697 Human Sap
37	80	15.4	80	7	ABU62252 Springoli
38	80	15.4	592	4	AAU05698 Human glu
39	68	13.1	85	4	AAH31912 Amino aci
40	66	12.7	83	4	AAH31929 Amino aci
41	61	11.8	61	2	AAW18584 Universal
42	61	11.8	69	2	AAV33258 Human pro
43	40	7.7	40	7	ABU62249 Springoli
44	38	7.3	38	7	ABU62250 Springoli
45	25	4.8	25	4	AAH67289 Human sap

ALIGNMENTS

RESULT 1

AAR70783

ID AAR70783 standard; protein; 524 AA.

AC AAR70783;

DT 25-MAR-2003 (revised)

DT 30-AUG-1995 (first entry)

XX Prosaposin.

KW Saposin-C; neuron; myelination; nervous system; neuroblastoma;  
KW neurotrophic peptide; multiple sclerosis; leukoencephalitis;  
KW adrenal leukodystrophy; prosaposin.

XX Homo sapiens.

OS WO9503821-A1.

PN 09-FEB-1995.

XX 28-JUL-1994; 94WO-US008453.

XX 30-JUL-1993; 93US-00100247.

XX 21-APR-1994; 94US-00232513.

XX (OBRI/) O'BRIEN J S.

XX O'brien JS, Kishimoto Y;

XX WPI; 1995-082029/11.

XX N-PSDB; AAQ85355.

XX Stimulating neural cell out-growth and myelination - with pro:saposin,  
XX saposin C or new neurotrophic peptide(s) from cytokine(s), for treating  
XX nervous system diseases.

XX Disclosure; Page 30-32; 50pp; English.

XX The peptide given in AAR70773, corresponding to amino acids 8-29 of human  
XX saposin-C (AAR70784), promotes neurite outgrowth in vitro. A consensus  
XX sequence was determined by comparing the peptide with hematopoietic and  
XX neurotrophic cytokines, and neurotrophic peptides (AAR70774-82) were  
XX identified in the AB loop of human ciliary neurotrophic factor,  
XX interleukin-6, -2, -3 and -gamma, erythropoietin and leukocyte  
XX inhibitory factor, and in helix C of human interleukin-1-beta and  
XX oncostatin-M. Prosaposin (AAR70783) and saposin-C also promoted nerve  
XX cell myelination ex vivo. (Updated on 25-MAR-2003 to correct PN field.)



```
CC (Updated on 25-MAR-2003 to correct PA field.) (Updated on 25-MAR-2003 to
CC correct PI field.)
XX
XX
SQ Sequence 524 AA;

Query Match      100.0%; Score 518; DB 2; Length 524;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 518; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LASLLGALAGPVLGLKECTRGSAVWQNVKTASDCGAVKHCLQTVWNNKPTVKSPLPCDIC 60
DB 7 LASLLGALAGPVLGLKECTRGSAVWQNVKTASDCGAVKHCLQTVWNNKPTVKSPLPCDIC 66
QY 61 KDVVTAAGDMLKDNATEEELVLEKTCDLWPKPNMSASCKEIVDSYLPVLDIIKGEMS 120
DB 67 KDVVTAAGDMLKDNATEEELVLEKTCDLWPKPNMSASCKEIVDSYLPVLDIIKGEMS 126
QY 121 RPGEVCSALNLCESLQKHLAELNHQKLESNKIPELDMTEVVAPFMANIPLLLYPDQDGR 180
DB 127 RPGEVCSALNLCESLQKHLAELNHQKLESNKIPELDMTEVVAPFMANIPLLLYPDQDGR 186
QY 181 SKPQKNDGVDQCDCIQMTVDIQTAVRTNSTFVQALVHVHKECDRLGPGMADICKNYS 240
DB 187 SKPQKNDGVDQCDCIQMTVDIQTAVRTNSTFVQALVHVHKECDRLGPGMADICKNYS 246
QY 241 QYSEIAIQMMHMQPKIEICALVGFCDVEKEMPQTLVPAKVASKNVIPALELVEPIKKE 300
DB 247 QYSEIAIQMMHMQPKIEICALVGFCDVEKEMPQTLVPAKVASKNVIPALELVEPIKKE 306
QY 301 VPAKSDVYCEVEFVLKVEVTKLIDNNKTEKELIDAFDQKCSKLPKLSJEEQEVVDVYGS 360
DB 307 VPAKSDVYCEVEFVLKVEVTKLIDNNKTEKELIDAFDQKCSKLPKLSJEEQEVVDVYGS 366
QY 361 SILSILLEVSPVLCVSMHLHCSGTRLPALTIVHTVTPQKDGGFCEVCKLVGYLDRLNLEKN 420
DB 367 SILSILLEVSPVLCVSMHLHCSGTRLPALTIVHTVTPQKDGGFCEVCKLVGYLDRLNLEKN 426
QY 421 STKQIBLALEKGCFLPDPYQKQCDQFVAYEPVLEILVENVDPFVCLKIGACPSAH 480
DB 427 STKQIBLALEKGCFLPDPYQKQCDQFVAYEPVLEILVENVDPFVCLKIGACPSAH 486
QY 481 KPLLGTCKIWPSPVWQNTETAQCNVAHECKRHVWN 518
DB 487 KPLLGTCKIWPSPVWQNTETAQCNVAHECKRHVWN 524

RESULT 2
AAW85652
ID AAW85652 standard; protein; 524 AA.
XX
XX AAW85652;
XX
XX
XX 19-JUL-1999 (first entry)
XX
XX Human prosaposin N-terminal peptide.
XX
XX Prosaposin; saposin; prosaptides; prosaposin receptor agonists; PRA;
XX peripheral nervous system; central nervous system; PNS; CNS; Akt; Bcl-2;
XX therapy; treatment; apoptosis; caspase; tumour necrosis factor; TNF;
XX cytokine; interferon gamma; IFN; inflammation; rheumatoid arthritis;
XX Crohn's disease; irritable bowel syndrome; asthma; cardiac infarction;
XX congestive heart failure; multiple sclerosis;
XX acute disseminated inflammatory leukoencephalitis;
XX progressive multifocal leukoencephalitis;
XX Parkinson's disease; amyotrophic lateral sclerosis; Huntington's disease;
XX ischemic heart disease; Guillain-Barre disease; alopecia; AIDS dementia;
XX cerebral malaria; HTLV; neuropathy;
XX inflammatory neurodegenerative disease; toxin-induced liver disease.
XX
XX Homo sapiens.
XX
XX OS
XX
XX
XX
XX W09912559-A1.
XX
```

```
PD 18-MAR-1999.
XX
XX PF 09-SEP-1998; 98WO-US019216.
XX
XX PR 09-SEP-1997; 97US-0058352P.
XX PR 04-JUN-1998; 98US-0088129P.
XX PA (REGC ) UNIV CALIFORNIA.
XX
XX PI O'brien JS;
XX
XX DR WPI; 1999-229139/19.
XX DR N-PSDB; AAX08488.
XX
XX PT Use of prosaposin receptor agonist.
XX
XX PS Claim 7; Fig 2; 90pp; English.
XX
XX CC Prosaposin is a 70kDa glycoprotein which is proteolytically processed to
XX generate saposins A, B, C and D, all of which are similar to each other
XX and have a similar placement of six cysteines, a glycosylation site and
XX conserved proline residues. Prosaposin, saposin C and prosaposin derived
XX peptides (prosaptides) have therapeutic applications in promoting
XX recovery after toxic, traumatic, myocardial ischaemic, degenerative and
XX inherited lesions to the peripheral and central nervous system.
XX Prosaposin receptor agonists (PRAs) inhibit proinflammatory cytokine-
XX induced apoptosis by activation of the Ser/Thr protein kinase Akt. Akt
XX dissociates complexes of Bcl-2 family members, such as BAD-Bcl-2,
XX releasing Bcl-2 and its family members which inhibit caspases, thereby
XX inhibiting apoptosis. An additional mechanism whereby PRAs inhibit
XX apoptosis is by blocking activation of JNK, a proapoptotic signaling
XX component. Within several minutes after binding to the receptor, PRAs
XX block JNK activation induced by tumor necrosis factor-alpha (TNF alpha).
XX The activation of JNK by TNF alpha is another well known mechanism for
XX TNF alpha-induced, as well as other proinflammatory cytokine-induced
XX apoptosis. The method can be used for inhibiting apoptosis which is
XX caspase-mediated or induced by a proinflammatory cytokine, for example
XX TNF alpha or interferon-gamma. It can be used for inhibiting apoptosis
XX associated with a disorder such as e.g. rheumatoid arthritis, Crohn's
XX disease, irritable bowel syndrome, asthma, cardiac infarction, congestive
XX heart failure, multiple sclerosis, acute disseminated inflammatory
XX leukoencephalitis, progressive multifocal leukoencephalitis, Alzheimer's
XX disease, Parkinson's disease, amyotrophic lateral sclerosis, Huntington's
XX disease, ischemic heart disease, Guillain-Barre disease, traumatic brain
XX injury, traumatic spinal cord injury, alopecia, AIDS dementia, cerebral
XX malaria, HTLV, neuropathy, inflammatory neurodegenerative disease, and
XX toxin-induced liver disease. This 524 N-terminal peptide of prosaposin
XX also acts as a prosaposin receptor agonist
XX
XX SQ Sequence 524 AA;

Query Match      100.0%; Score 518; DB 2; Length 524;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 518; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LASLLGALAGPVLGLKECTRGSAVWQNVKTASDCGAVKHCLQTVWNNKPTVKSPLPCDIC 60
DB 7 LASLLGALAGPVLGLKECTRGSAVWQNVKTASDCGAVKHCLQTVWNNKPTVKSPLPCDIC 66
QY 61 KDVVTAAGDMLKDNATEEELVLEKTCDLWPKPNMSASCKEIVDSYLPVLDIIKGEMS 120
DB 67 KDVVTAAGDMLKDNATEEELVLEKTCDLWPKPNMSASCKEIVDSYLPVLDIIKGEMS 126
QY 121 RPGEVCSALNLCESLQKHLAELNHQKLESNKIPELDMTEVVAPFMANIPLLLYPDQDGR 180
DB 127 RPGEVCSALNLCESLQKHLAELNHQKLESNKIPELDMTEVVAPFMANIPLLLYPDQDGR 186
QY 181 SKPQKNDGVDQCDCIQMTVDIQTAVRTNSTFVQALVHVHKECDRLGPGMADICKNYS 240
DB 187 SKPQKNDGVDQCDCIQMTVDIQTAVRTNSTFVQALVHVHKECDRLGPGMADICKNYS 246
QY 241 QYSEIAIQMMHMQPKIEICALVGFCDVEKEMPQTLVPAKVASKNVIPALELVEPIKKE 300
DB 247 QYSEIAIQMMHMQPKIEICALVGFCDVEKEMPQTLVPAKVASKNVIPALELVEPIKKE 306
QY 301 VPAKSDVYCEVEFVLKVEVTKLIDNNKTEKELIDAFDQKCSKLPKLSJEEQEVVDVYGS 360
DB 307 VPAKSDVYCEVEFVLKVEVTKLIDNNKTEKELIDAFDQKCSKLPKLSJEEQEVVDVYGS 366
QY 361 SILSILLEVSPVLCVSMHLHCSGTRLPALTIVHTVTPQKDGGFCEVCKLVGYLDRLNLEKN 420
DB 367 SILSILLEVSPVLCVSMHLHCSGTRLPALTIVHTVTPQKDGGFCEVCKLVGYLDRLNLEKN 426
QY 421 STKQIBLALEKGCFLPDPYQKQCDQFVAYEPVLEILVENVDPFVCLKIGACPSAH 480
DB 427 STKQIBLALEKGCFLPDPYQKQCDQFVAYEPVLEILVENVDPFVCLKIGACPSAH 486
QY 481 KPLLGTCKIWPSPVWQNTETAQCNVAHECKRHVWN 518
DB 487 KPLLGTCKIWPSPVWQNTETAQCNVAHECKRHVWN 524

RESULT 2
AAW85652
ID AAW85652 standard; protein; 524 AA.
XX
XX AAW85652;
XX
XX
XX 19-JUL-1999 (first entry)
XX
XX Human prosaposin N-terminal peptide.
XX
XX Prosaposin; saposin; prosaptides; prosaposin receptor agonists; PRA;
XX peripheral nervous system; central nervous system; PNS; CNS; Akt; Bcl-2;
XX therapy; treatment; apoptosis; caspase; tumour necrosis factor; TNF;
XX cytokine; interferon gamma; IFN; inflammation; rheumatoid arthritis;
XX Crohn's disease; irritable bowel syndrome; asthma; cardiac infarction;
XX congestive heart failure; multiple sclerosis;
XX acute disseminated inflammatory leukoencephalitis;
XX progressive multifocal leukoencephalitis;
XX Parkinson's disease; amyotrophic lateral sclerosis; Huntington's disease;
XX ischemic heart disease; Guillain-Barre disease; alopecia; AIDS dementia;
XX cerebral malaria; HTLV; neuropathy;
XX inflammatory neurodegenerative disease; toxin-induced liver disease.
XX
XX Homo sapiens.
XX
XX OS
XX
XX
XX
XX W09912559-A1.
XX
```



Db 247 QYSEIAIOMMHQPKKEICALVGFCDCEVKEMPMQTLVPAKVASKNVIPALVELPEIKKHE 306  
Qy 301 VPAKSDVYCEFLVKEVTKLIDNNKTEKEILDAFKMCSKLPKSLSESCQEVVDITYGS 360  
Db 307 VPAKSDVYCEFLVKEVTKLIDNNKTEKEILDAFKMCSKLPKSLSESCQEVVDITYGS 366  
Qy 361 SILSILLEEVSPELVCSMLHLCSTRLPALTVHVTQPKDGGFCVCKLVGYLDRLNLEKN 420  
Db 367 SILSILLEEVSPELVCSMLHLCSTRLPALTVHVTQPKDGGFCVCKLVGYLDRLNLEKN 426  
Qy 421 STKQEIILAALEKGCSEFLPDYPKQCDQFVAEYEPVLEILVEVMDPSFVCLKIGACPSAH 480  
Db 427 STKQEIILAALEKGCSEFLPDYPKQCDQFVAEYEPVLEILVEVMDPSFVCLKIGACPSAH 486  
Qy 481 KPLLGTEKCIWGPSYWCNTETAAQCNVAHECKRHVWN 518  
Db 487 KPLLGTEKCIWGPSYWCNTETAAQCNVAHECKRHVWN 524

RESULT 3  
AA58716  
ID AA58716 standard; protein; 524 AA.  
AC AA58716;  
XX  
DT 25-APR-2000 (first entry)  
XX Human prosaposin.  
XX  
KW Prosaposin; saponin B; antiangiogenic; angiogenesis inhibitor;  
KW antitumor; antiproliferative; antimigratory; Kaposi's sarcoma; tumour;  
KW human; therapy.  
XX  
OS Homo sapiens.  
XX  
XX Key Location/Qualifiers  
FH Protein 195..275  
FT /note= "mature saposin B"  
FT Peptide 195..205  
FT /note= "specifically claimed antiangiogenic peptide of  
FT Claim 23"  
FT Peptide 196..200  
FT /note= "specifically claimed antiangiogenic peptide of  
FT Claim 4"  
XX WO200002902-A1.  
XX  
XX 20-JAN-2000.  
XX  
XX 12-JUL-1999; 99WO-US015772.  
XX  
XX 13-JUL-1998; 98US-0092647P.  
XX  
XX (GILL/) GILL P S.  
XX  
XX Gill PS;  
XX  
XX WPI; 2000-171128/15.  
XX  
XX Saposin B derived peptides, useful as inhibitors of angiogenesis and  
XX tumor growth.  
XX  
XX Disclosure; Page 18; 78pp; English.  
XX  
XX The present sequence is that of human prosaposin, a precursor of saposin  
XX B. The invention is based on the discovery that saposin B, previously  
XX known to be involved in the hydrolysis of sphingolipids, has potent  
XX antiangiogenic and antitumor activity, and also has antiproliferative  
XX and antimigratory activity against endothelial cells. This activity is  
XX conserved in cryptic polypeptides as small as 5 amino acids (see AAY58684  
XX -715), which can be synthetically prepared and used in vitro or in vivo  
XX for the treatment of undesired angiogenesis and tumor growth, especially  
XX Kaposi's sarcoma (claimed). The polypeptides can also be used in

CC conjunction with cytotoxic moieties to selectively kill certain cell  
CC types, e.g. for treatment of cancer, angiofibroma, neovascular glaucoma,  
CC arteriovenous malformation, nonunion fracture, arthritis and other  
CC connective tissue disorders, Osler-Weber syndrome, atherosclerotic  
CC plaque, psoriasis, corneal graft neovascularization, pyogenic granuloma,  
CC retrolental fibroplasia, diabetic retinopathy, scleroderma, haemangioma,  
CC trachoma, vascular adhesions and hypertrophic scars  
XX  
SQ Sequence 524 AA;  
Query Match 100.0%; Score 518; DB 3; Length 524;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 518; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 1 LASLGAALAGPVGLKECTRGSAVMQNVKTASDCGAVKHCLQTVNMKPTVKSPLPCDIC 60  
Db 7 LASLGAALAGPVGLKECTRGSAVMQNVKTASDCGAVKHCLQTVNMKPTVKSPLPCDIC 66  
Qy 61 KDVTAAAGDMLKDNATEEELVLEKTCDWLKPNNMSASCKEIVDSYLPVLDIIKEMS 120  
Db 67 KDVTAAAGDMLKDNATEEELVLEKTCDWLKPNNMSASCKEIVDSYLPVLDIIKEMS 126  
Qy 121 RPGEVCSALNLCESLOKHLAELNHQKLESNKIPELDMTEVAVPFMANIPLLLYPQDQPR 180  
Db 127 RPGEVCSALNLCESLOKHLAELNHQKLESNKIPELDMTEVAVPFMANIPLLLYPQDQPR 186  
Qy 181 SKPQKNDGVDVDCDCIQMVTDIQTAVRTNSTFTVQALVEHVKECDRLGPGMADICKNYIS 240  
Db 187 SKPQKNDGVDVDCDCIQMVTDIQTAVRTNSTFTVQALVEHVKECDRLGPGMADICKNYIS 246  
Qy 241 QYSEIAIOMMHQPKKEICALVGFCDCEVKEMPMQTLVPAKVASKNVIPALVELPEIKKHE 300  
Db 247 QYSEIAIOMMHQPKKEICALVGFCDCEVKEMPMQTLVPAKVASKNVIPALVELPEIKKHE 306  
Qy 301 VPAKSDVYCEFLVKEVTKLIDNNKTEKEILDAFKMCSKLPKSLSESCQEVVDITYGS 360  
Db 307 VPAKSDVYCEFLVKEVTKLIDNNKTEKEILDAFKMCSKLPKSLSESCQEVVDITYGS 366  
Qy 361 SILSILLEEVSPELVCSMLHLCSTRLPALTVHVTQPKDGGFCVCKLVGYLDRLNLEKN 420  
Db 367 SILSILLEEVSPELVCSMLHLCSTRLPALTVHVTQPKDGGFCVCKLVGYLDRLNLEKN 426  
Qy 421 STKQEIILAALEKGCSEFLPDYPKQCDQFVAEYEPVLEILVEVMDPSFVCLKIGACPSAH 480  
Db 427 STKQEIILAALEKGCSEFLPDYPKQCDQFVAEYEPVLEILVEVMDPSFVCLKIGACPSAH 486  
Qy 481 KPLLGTEKCIWGPSYWCNTETAAQCNVAHECKRHVWN 518  
Db 487 KPLLGTEKCIWGPSYWCNTETAAQCNVAHECKRHVWN 524

RESULT 4  
ABU79099  
ID ABU79099 standard; protein; 524 AA.  
XX AC  
XX ABU79099;  
XX  
XX 18-JUN-2003 (first entry)  
XX  
XX Lip-TMA binding protein, Prosaposin.  
XX  
XX Superantigen; SAG; staphylococcal enterotoxin; tumour; cancer; apoptosis;  
XX gene therapy; mammalian cell receptor; cytostatic;  
XX tumour associated lipid; anergy; T cell; antigen presenting cell; APC;  
XX tumouricidal immunocyte; antitumour.  
XX Unidentified.  
XX US2002177551-A1.  
XX  
XX 28-NOV-2002.  
XX  
XX 30-MAY-2001; 2001US-00870759.  
XX PF

```
XX PR 31-MAY-2000; 2000US-0208128P.
XX PA (TERM/) TERMAN D S.
XX PI Terman DS;
XX DR WPI; 2003-361759/34.
XX XX
XX PT A mammalian cell receptor, useful in the treatment of cancer by binding
XX PT to tumor associated lipids where the binding induces anergy or apoptosis
XX PT in T cells and antigen presenting cells.
XX PS Disclosure; Page; 167pp; English.
XX XX
XX CC The invention relates to a mammalian cell receptor, useful in the
XX CC treatment of cancer, which binds to tumor associated lipids and induces
XX CC anergy or apoptosis in the T cells and antigen presenting cells (APCs).
XX CC Also included are a mammalian cell useful in the treatment of cancer
XX CC where the receptor which binds tumor associated lipids and induces
XX CC cellular inactivation or death is deleted or functionally deactivated,
XX CC producing (M1) a tumouricidal immunocyte population in vivo in a mammal
XX CC (by allowing tumour associated lipids to contact immunocytes in which
XX CC receptors for immunosuppressive fatty acids, ceramides, glycolipids,
XX CC sphingolipids, glycosphingolipids, phosphosphingolipids, gangliosides,
XX CC sialylated glycans, lipopeptides and proteoglycolipids are inactivated or
XX CC deleted), a construct useful in the treatment of cancer comprising a
XX CC superantigen (SAG) nucleotide inserted into a virus, a mammalian T cell
XX CC useful in the treatment of cancer (where an adaptor protein which
XX CC inhibits T cell activation by tumour associated antigens is deleted or
XX CC functionally deactivated), a composition useful in the treatment of
XX CC cancer (comprising a lipid raft conjugated to a superantigen), producing
XX CC (M2) a tumouricidal immunocyte population ex vivo in a mammal (by
XX CC allowing tumour associated lipids to contact immunocytes, in which
XX CC receptors for the lipids are inactivated or deleted to produce a
XX CC tumouricidal immunocyte population, and administering the tumouricidal
XX CC activated immunocytes to the host), producing (M3) a tumouricidal APC
XX CC population ex vivo in a mammal (by allowing a tumour associated lipid to
XX CC contact APCs, in which receptors for the tumour associated lipids are
XX CC inactivated or deleted to produce a tumouricidal activated lipids are
XX CC and administering APCs to the host), producing a tumouricidal T cell
XX CC population ex vivo in a mammal (by allowing a tumour associated lipid to
XX CC contact T cells, in which adaptor proteins, which inhibit T cell
XX CC activation by tumour associated antigens, are deleted or functionally
XX CC deactivated to produce a tumouricidal population of T cells, and
XX CC administering the tumouricidal activated T cells to the host, or
XX CC allowing a superantigen-lipid raft to contact T cells ex vivo, and
XX CC administering the tumouricidal activated T cells to the host), treating
XX CC (M5) cancer in a mammal (by administering a lipid binding molecule which
XX CC binds immunosuppressive tumour associated lipids in vivo), producing (M6)
XX CC a tumouricidal T cell population in vivo in a mammal (by allowing a
XX CC tumour associated antigen to contact immunocytes in which adaptor
XX CC proteins which inhibit T cell activation by tumour associated antigens
XX CC are deleted or functionally deactivated) and producing (M7) a
XX CC tumouricidal T cell population ex vivo in a mammal comprising allowing a
XX CC superantigen-lipid raft conjugate to contact immunocytes in vivo. The
XX CC receptors, methods and compositions are useful for treating cancers and
XX CC tumours. Bacterial superantigens are co-administered or administered as
XX CC fusion constructs with anti-tumour proteins or motifs. The present
XX CC sequence represents a tumour antigen or a motif identifying a tumour
XX CC antigen, which can be functionally deactivated in the method of the
XX CC invention. Note: The sequence data for this patent did not form part of
XX CC the printed specification, but was obtained in electronic format from the
XX CC US patent office website at
XX CC "seqdata.uspto.gov/sequence.html?DocID=20020177551"
XX CC
XX CC Sequence 524 AA;
XX CC
XX CC Query Match 100.0%; Score 518; DB 6; Length 524;
XX CC Best Local Similarity 100.0%; Pred. No. 0;
XX CC Matches 518; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX CC
XX CC 1 LASLGAALAGPVLGLKECTRGSAVWQNVKTASDCGAVKHCLQTVWKNKPTVKSPLPCDIC 60
XX CC
```

PT New polypeptides (e.g. kinases, phosphatases, proteases, transporters, cytoskeletal proteins, receptors or transcription factors), useful for treating cancer, e.g. colon cancer, gastric cancer, sarcoma, lymphoma or leukemia.

XX Example 2; SEQ ID NO 1866; 134pp; English.

PS The invention describes a purified polypeptide, which comprises a fragment of a kinase, phosphatase, protease, protease inhibitor, transporter, cytoskeletal protein, receptor or transcription factor. The polypeptide is useful as an immunogenic composition for eliciting in a mammal an immunogenic response directed against any of the purified polypeptide. The purified polypeptide, or the antibody that binds to this polypeptide, is useful for treating cancer. The polypeptide is also useful for identifying compounds that binds to a naturally processed class I or class II MHC-binding polypeptide. The polypeptides and polynucleotides are particularly useful for treating or preventing myeloma, colon cancer, gastric cancer, adenocarcinoma, sarcoma, melanoma, lymphoma or leukaemia. These are also useful for screening agents for treating the above mentioned diseases. This sequence represents an expressed protein tag (SPT) isolated from human tissue for translational profiling. Note: This sequence does not appear in the printed specification but was obtained in electronic format directly from WIPO at ftp.wipo.int/pub/published\_pct\_sequences

XX Sequence 524 AA;

Query Match 100.0%; Score 518; DB 6; Length 524;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 518; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LASLGAALAGPVLGLKECTRGSAVWCNQVKTASDCGAVKHCLQTVWVKPTVKSLPCDIC 60  
Db 7 LASLGAALAGPVLGLKECTRGSAVWCNQVKTASDCGAVKHCLQTVWVKPTVKSLPCDIC 66

QY 61 KDVTAAAGDMLKDNATEEILVYLEKTCMDLPKNMSASCKEIVDSYLPVLDIIGKEMS 120  
Db 67 KDVTAAAGDMLKDNATEEILVYLEKTCMDLPKNMSASCKEIVDSYLPVLDIIGKEMS 126

QY 121 RPGEVCSALNLCESIQKHLAEINHQKLESNKIPELDMTEVVAPPMANIPLLLPDQGP 180  
Db 127 RPGEVCSALNLCESIQKHLAEINHQKLESNKIPELDMTEVVAPPMANIPLLLPDQGP 186

QY 181 SKPQKNDGVQCDCIQMTDITQAVRTNSTFVQALVEHVKEECDRLGPGMADICKNYIS 240  
Db 187 SKPQKNDGVQCDCIQMTDITQAVRTNSTFVQALVEHVKEECDRLGPGMADICKNYIS 246

QY 241 QYSEIAIQMMHMOPKEICALVGFCDEVKPMQTLVPAKVASKNVIPAELVPEIKHE 300  
Db 247 QYSEIAIQMMHMOPKEICALVGFCDEVKPMQTLVPAKVASKNVIPAELVPEIKHE 306

QY 301 VPAKSDVYCEVCEFLVKEVTKLIDNNKTEKILDAFDKMSKLPKSLSEECQEVVDITYGS 360  
Db 307 VPAKSDVYCEVCEFLVKEVTKLIDNNKTEKILDAFDKMSKLPKSLSEECQEVVDITYGS 366

QY 361 SILSILLEEVSPVLCVSMHLCSGTRLPALTVHTVTPQKDGFCVCKLGVYLDRLNLEKN 420  
Db 367 SILSILLEEVSPVLCVSMHLCSGTRLPALTVHTVTPQKDGFCVCKLGVYLDRLNLEKN 426

QY 421 STKQEILAALEKSGCSFLDPYQKCDQFVAEYEPVLILVEVMDPSPVCLIKGACPSAH 480  
Db 427 STKQEILAALEKSGCSFLDPYQKCDQFVAEYEPVLILVEVMDPSPVCLIKGACPSAH 486

QY 481 KPLLGTCKIWPSPYWCQNTETAACNAVEHCKRHVN 518  
Db 487 KPLLGTCKIWPSPYWCQNTETAACNAVEHCKRHVN 524

RESULT 6  
ID ABU05207  
XX ABU05207 standard; protein; 524 AA.  
AC ABU05207;

XX 29-JAN-2003 (first entry)  
DT Human expressed protein tag (EPT) #1873.  
DE  
XX

KW Translational profiling; expressed protein tag; EPT; kinase; phosphatase; protease; protease inhibitor; transporter; cytoskeletal protein; receptor; transcription factor; cancer; MHC;  
KW major histocompatibility complex; myeloma; colon cancer; gastric cancer; adenocarcinoma; sarcoma; melanoma; lymphoma; leukaemia.

OS Homo sapiens.  
XX WO200278524-A2.  
XX 10-OCT-2002.

XX 28-MAR-2002; 2002WO-US009671.  
XX 28-MAR-2001; 2001US-0279495P.  
PR 21-MAY-2001; 2001US-0292544P.  
PR 08-AUG-2001; 2001US-0310801P.  
PR 01-OCT-2001; 2001US-0326370P.  
PR 04-DEC-2001; 2001US-0336780P.  
PR 20-FEB-2002; 2002US-0358985P.

XX (ZYCO-) ZYCOS INC.

XX Chicx RM, Tomlinson AJ, Urban RG;  
XX WPI; 2003-040607/03.

XX New polypeptides (e.g. kinases, phosphatases, proteases, transporters, cytoskeletal proteins, receptors or transcription factors), useful for treating cancer, e.g. colon cancer, gastric cancer, sarcoma, lymphoma or leukemia.

PS Example 2; SEQ ID NO 1873; 134pp; English.

XX The invention describes a purified polypeptide, which comprises a fragment of a kinase, phosphatase, protease, protease inhibitor, transporter, cytoskeletal protein, receptor or transcription factor. The polypeptide is useful as an immunogenic composition for eliciting in a mammal an immunogenic response directed against any of the purified polypeptide. The purified polypeptide, or the antibody that binds to this polypeptide, is useful for treating cancer. The polypeptide is also useful for identifying compounds that binds to a naturally processed class I or class II MHC-binding polypeptide. The polypeptides and polynucleotides are particularly useful for treating or preventing myeloma, colon cancer, gastric cancer, adenocarcinoma, sarcoma, melanoma, lymphoma or leukaemia. These are also useful for screening agents for treating the above mentioned diseases. This sequence represents an expressed protein tag (EPT) isolated from human tissue for translational profiling. Note: This sequence does not appear in the printed specification but was obtained in electronic format directly from WIPO at ftp.wipo.int/pub/published\_pct\_sequences

XX Sequence 524 AA;

Query Match 100.0%; Score 518; DB 6; Length 524;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 518; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LASLGAALAGPVLGLKECTRGSAVWCNQVKTASDCGAVKHCLQTVWVKPTVKSLPCDIC 60  
Db 7 LASLGAALAGPVLGLKECTRGSAVWCNQVKTASDCGAVKHCLQTVWVKPTVKSLPCDIC 66

QY 61 KDVTAAAGDMLKDNATEEILVYLEKTCMDLPKNMSASCKEIVDSYLPVLDIIGKEMS 120  
Db 67 KDVTAAAGDMLKDNATEEILVYLEKTCMDLPKNMSASCKEIVDSYLPVLDIIGKEMS 126

QY 121 RPGEVCSALNLCESIQKHLAEINHQKLESNKIPELDMTEVVAPPMANIPLLLPDQGP 180  
Db 127 RPGEVCSALNLCESIQKHLAEINHQKLESNKIPELDMTEVVAPPMANIPLLLPDQGP 186

Db 127 RPEVCSALNLCESLQKHLAEINHQKLESNKIPELDMTEVWAPPFMANIPILLYPQDGR 186  
 QY 181 SKPQPKDNGDVQCDCIQMVTDTQTAVRTNSTFVQALVEHVKEECDRLGPGMADICKNYIS 240  
 Db 187 SKPQPKDNGDVQCDCIQMVTDTQTAVRTNSTFVQALVEHVKEECDRLGPGMADICKNYIS 246  
 QY 241 QYSEIAIQMMHMQPKEICALVGFCDVEKMPQTLVPAKVASKNVTPALELVEPIKKE 300  
 Db 247 QYSEIAIQMMHMQPKEICALVGFCDVEKMPQTLVPAKVASKNVTPALELVEPIKKE 306  
 QY 301 VPAKSDVYCEVCEFLVKEVTKLIDNNKTEKILDAFDMCKSLPKSLSECEQVVDITYGS 360  
 Db 307 VPAKSDVYCEVCEFLVKEVTKLIDNNKTEKILDAFDMCKSLPKSLSECEQVVDITYGS 366  
 QY 361 SILSILLEEVSPELVCSMLHLCGSTRLPALTHTVHTQPKDGFCEVCKLVGLVLDRLNLEKN 420  
 Db 367 SILSILLEEVSPELVCSMLHLCGSTRLPALTHTVHTQPKDGFCEVCKLVGLVLDRLNLEKN 426  
 QY 421 STKQEIILAALEKGCSEFLPDYPYQKCDQFVAEYEPVLIIEILVEVMDPSFVCLKIGACPSAH 480  
 Db 427 STKQEIILAALEKGCSEFLPDYPYQKCDQFVAEYEPVLIIEILVEVMDPSFVCLKIGACPSAH 486  
 QY 481 KPLLGTCKICWGPSYWCQNTETAACNAVEHCKRHVWN 518  
 Db 487 KPLLGTCKICWGPSYWCQNTETAACNAVEHCKRHVWN 524

RESULT 7  
 ABU05203  
 ID ABU05203 standard; protein; 524 AA.  
 XX  
 AC ABU05203;  
 XX  
 DT 29-JAN-2003 (first entry)  
 XX  
 DE Human expressed protein tag (EPT) #1869.  
 XX  
 KW Translational profiling; expressed protein tag; EPT; kinase; phosphatase;  
 KW protease; protease inhibitor; transporter; cytoskeletal protein;  
 KW receptor; transcription factor; cancer; MHC;  
 KW major histocompatibility complex; myeloma; colon cancer; gastric cancer;  
 KW adenocarcinoma; sarcoma; melanoma; lymphoma; leukaemia.  
 XX  
 OS Homo sapiens.  
 XX  
 PN WO200278524-A2.  
 XX  
 PD 10-OCT-2002.  
 XX  
 PF 28-MAR-2002; 2002WO-US009671.  
 XX  
 PR 28-MAR-2001; 2001US-0279495P.  
 PR 21-MAY-2001; 2001US-0292544P.  
 PR 08-AUG-2001; 2001US-0310801P.  
 PR 01-OCT-2001; 2001US-0326370P.  
 PR 04-DEC-2001; 2001US-0336780P.  
 PR 20-FEB-2002; 2002US-0358985P.  
 XX  
 PA (ZYCO-) ZYCOS INC.  
 XX  
 PI Chicx RM, Tomlinson AJ, Urban RG;  
 XX  
 DR WPI; 2003-040607/03.  
 XX  
 PT New polypeptides (e.g. kinases, phosphatases, proteases, transporters,  
 PT cytoskeletal proteins, receptors or transcription factors), useful for  
 PT treating cancer, e.g. colon cancer, gastric cancer, sarcoma, lymphoma or  
 PT leukemia.  
 XX  
 PS Example 2; SEQ ID NO 1869; 134pp; English.  
 XX  
 CC The invention describes a purified polypeptide, which comprises a  
 CC fragment of a kinase, phosphatase, protease, protease inhibitor,

CC transporter, cytoskeletal protein, receptor or transcription factor. The  
 CC polypeptide is useful as an immunogenic composition for eliciting in a  
 CC mammal an immunogenic response directed against any of the purified  
 CC polypeptide. The purified polypeptide, or the antibody that binds to this  
 CC polypeptide, is useful for treating cancer. The polypeptide is also  
 CC useful for identifying compounds that binds to a naturally processed  
 CC class I or class II MHC-binding polypeptide. The polypeptides and  
 CC polynucleotides are particularly useful for treating or preventing  
 CC myeloma, colon cancer, gastric cancer, adenocarcinoma, sarcoma, melanoma,  
 CC lymphoma or leukaemia. These are also useful for screening agents for  
 CC treating the above mentioned diseases. This sequence represents an  
 CC expressed protein tag (EPT) isolated from human tissue for translational  
 CC profiling. Note: This sequence does not appear in the printed  
 CC specification but was obtained in electronic format directly from WIPO at  
 CC ftp.wipo.int/pub/published\_pct\_sequences  
 XX  
 SQ Sequence 524 AA;

Query Match 100.0%; Score 518; DB 6; Length 524;  
 Best Local Similarity 100.0%; Pred. No. 0;  
 Matches 518; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 LASLGAALAGPVGLGKECTRGSAVWCQNVKTASDCGAVKHCLQTVMNKPTVKSPLPCDIC 60  
 Db 7 LASLGAALAGPVGLGKECTRGSAVWCQNVKTASDCGAVKHCLQTVMNKPTVKSPLPCDIC 66  
 QY 61 KDVTVAAGDMLKDNATEEELVYLEKTCDWLPKPNMSASCKEIVDSYLPVILDIITKEMS 120  
 Db 67 KDVTVAAGDMLKDNATEEELVYLEKTCDWLPKPNMSASCKEIVDSYLPVILDIITKEMS 126  
 QY 121 RPEVCSALNLCESLQKHLAEINHQKLESNKIPELDMTEVWAPPFMANIPILLYPQDGR 180  
 Db 127 RPEVCSALNLCESLQKHLAEINHQKLESNKIPELDMTEVWAPPFMANIPILLYPQDGR 186  
 QY 181 SKPQPKDNGDVQCDCIQMVTDTQTAVRTNSTFVQALVEHVKEECDRLGPGMADICKNYIS 240  
 Db 187 SKPQPKDNGDVQCDCIQMVTDTQTAVRTNSTFVQALVEHVKEECDRLGPGMADICKNYIS 246  
 QY 241 QYSEIAIQMMHMQPKEICALVGFCDVEKMPQTLVPAKVASKNVTPALELVEPIKKE 300  
 Db 247 QYSEIAIQMMHMQPKEICALVGFCDVEKMPQTLVPAKVASKNVTPALELVEPIKKE 306  
 QY 301 VPAKSDVYCEVCEFLVKEVTKLIDNNKTEKILDAFDMCKSLPKSLSECEQVVDITYGS 360  
 Db 307 VPAKSDVYCEVCEFLVKEVTKLIDNNKTEKILDAFDMCKSLPKSLSECEQVVDITYGS 366  
 QY 361 SILSILLEEVSPELVCSMLHLCGSTRLPALTHTVHTQPKDGFCEVCKLVGLVLDRLNLEKN 420  
 Db 367 SILSILLEEVSPELVCSMLHLCGSTRLPALTHTVHTQPKDGFCEVCKLVGLVLDRLNLEKN 426  
 QY 421 STKQEIILAALEKGCSEFLPDYPYQKCDQFVAEYEPVLIIEILVEVMDPSFVCLKIGACPSAH 480  
 Db 427 STKQEIILAALEKGCSEFLPDYPYQKCDQFVAEYEPVLIIEILVEVMDPSFVCLKIGACPSAH 486  
 QY 481 KPLLGTCKICWGPSYWCQNTETAACNAVEHCKRHVWN 518  
 Db 487 KPLLGTCKICWGPSYWCQNTETAACNAVEHCKRHVWN 524

RESULT 8  
 ABU07340  
 ID ABU07340 standard; protein; 524 AA.  
 XX  
 AC ABU07340;  
 XX  
 DT 29-JAN-2003 (first entry)  
 XX  
 DE Human expressed protein tag (EPT) #2041.  
 XX  
 KW Translational profiling; expressed protein tag; EPT; kinase; phosphatase;  
 KW protease; protease inhibitor; transporter; cytoskeletal protein;  
 KW receptor; transcription factor; cancer; MHC;  
 KW major histocompatibility complex; myeloma; colon cancer; gastric cancer;

adenocarcinoma; sarcoma; melanoma; lymphoma; leukaemia.

Homo sapiens.

WO200278524-A2.

10-OCT-2002.

28-MAR-2002; 2002WO-US009671.

28-MAR-2001; 2001US-0279495P.

21-MAY-2001; 2001US-0292544P.

08-AUG-2001; 2001US-0310801P.

01-OCT-2001; 2001US-0326370P.

04-DEC-2001; 2001US-0336780P.

20-FEB-2002; 2002US-0358985P.

(ZYCO-) ZYCOS INC.

Chicz RM, Tomlinson AJ, Urban RG;

WPI; 2003-040607/03.

New polypeptides (e.g. kinases, phosphatases, proteases, transporters, cytoskeletal proteins, receptors or transcription factors), useful for treating cancer, e.g. colon cancer, gastric cancer, sarcoma, lymphoma or leukemia.

Example 2; SEQ ID NO 2041; 134pp; English.

The invention describes a purified polypeptide, which comprises a fragment of a kinase, phosphatase, protease, protease inhibitor, transporter, cytoskeletal protein, receptor or transcription factor. The polypeptide is useful as an immunogenic composition for eliciting in a mammal an immunogenic response directed against any of the purified polypeptide. The purified polypeptide, or the antibody that binds to this polypeptide. The purified polypeptide, or the antibody that binds to this polypeptide, is useful for treating cancer. The polypeptide is also useful for identifying compounds that binds to a naturally processed class I or class II MHC-binding polypeptide. The polypeptides and polynucleotides are particularly useful for treating or preventing myeloma, colon cancer, gastric cancer, adenocarcinoma, sarcoma, melanoma, lymphoma or leukaemia. These are also useful for screening agents for treating the above mentioned diseases. This sequence represents an expressed protein tag (EPT) isolated from human tissue for translational profiling. Note: This sequence does not appear in the printed specification but was obtained in electronic format directly from WIPO at ftp.wipo.int/pub/published\_pct\_sequences

Sequence 524 AA;

Query Match 100.0%; Score 518; DB 6; Length 524;

Best Local Similarity 100.0%; Pred. No. 0;

Matches 518; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 LASLLGALAGPVLGLKECTRGSAVWCONVTASDCGAVKHCLQTVWKNKPTVKSLPCDIC 60

Db 7 LASLLGALAGPVLGLKECTRGSAVWCONVTASDCGAVKHCLQTVWKNKPTVKSLPCDIC 66

Qy 61 KDVVTAAGDMLKDNATTEELIYVLEKTCDWLPKPNMSASCKEIVDSYLPVLDIIRKEMS 120

Db 67 KDVVTAAGDMLKDNATTEELIYVLEKTCDWLPKPNMSASCKEIVDSYLPVLDIIRKEMS 126

Qy 121 RPGEVCSALNLCESLQKHLAEINHQKLESNKIPELDMEVAVPFMANIPILLYPQDQPR 180

Db 127 RPGEVCSALNLCESLQKHLAEINHQKLESNKIPELDMEVAVPFMANIPILLYPQDQPR 186

Qy 181 SKQPQKNGDVQCQICQWVTDIQTAVRTNSTFVQALVEHVKECDRLGPGWADICKNYS 240

Db 187 SKQPQKNGDVQCQICQWVTDIQTAVRTNSTFVQALVEHVKECDRLGPGWADICKNYS 246

Qy 241 QYSEIATQMMHMQPKEICALVGCDCEVKNPMQTLVPKAVSKNVIPALELVEPIKKHE 300

Db 247 QYSEIATQMMHMQPKEICALVGCDCEVKNPMQTLVPKAVSKNVIPALELVEPIKKHE 306

Qy 301 VPAKSDVYCEVCFVLVKEVTKLIDNNKTEKEILDADFQKMSCKLPKSLSEECQEVVDITYGS 360

Db 307 VPAKSDVYCEVCFVLVKEVTKLIDNNKTEKEILDADFQKMSCKLPKSLSEECQEVVDITYGS 366

Qy 361 SILSILLEEVSPELVCSMLHLCSTRLPALTVHVTQPKDGGFCFVCKKLVGLDRNLEKN 420

Db 367 SILSILLEEVSPELVCSMLHLCSTRLPALTVHVTQPKDGGFCFVCKKLVGLDRNLEKN 426

Qy 421 STKQEIILAALKEKGSFLPDYQKOCDOFVAEYEPVLIEILVEVMDPSFVCLKIGACPSAH 480

Db 427 STKQEIILAALKEKGSFLPDYQKOCDOFVAEYEPVLIEILVEVMDPSFVCLKIGACPSAH 486

Qy 481 KPLLGTCKCIWGPSYWCQNTETATAACNAVEHCKRHVWN 518

Db 487 KPLLGTCKCIWGPSYWCQNTETATAACNAVEHCKRHVWN 524

RESULT 9

ABU05216

ID ABU05216 standard; protein; 524 AA.

XX XX

AC ABU05216;

XX XX

DT 29-JAN-2003 (first entry)

XX XX

DE Human expressed protein tag (EPT) #1882.

XX XX

KW Translational profiling; expressed protein tag; EPT; kinase; phosphatase; protease; protease inhibitor; transporter; cytoskeletal protein;

KW receptor; transcription factor; cancer; MHC;

KW major histocompatibility complex; myeloma; colon cancer; gastric cancer;

KW adenocarcinoma; sarcoma; melanoma; lymphoma; leukaemia.

XX XX

OS Homo sapiens.

XX XX

PN WO200278524-A2.

XX XX

PD 10-OCT-2002.

XX XX

PF 28-MAR-2002; 2002WO-US009671.

XX XX

PR 28-MAR-2001; 2001US-0279495P.

XX XX

PR 21-MAY-2001; 2001US-0292544P.

XX XX

PR 08-AUG-2001; 2001US-0310801P.

XX XX

PR 01-OCT-2001; 2001US-0326370P.

XX XX

PR 04-DEC-2001; 2001US-0336780P.

XX XX

PR 20-FEB-2002; 2002US-0358985P.

XX XX

PA (ZYCO-) ZYCOS INC.

XX XX

PI Chicx RM, Tomlinson AJ, Urban RG;

XX XX

DR WPI; 2003-040607/03.

XX XX

XX New polypeptides (e.g. kinases, phosphatases, proteases, transporters, cytoskeletal proteins, receptors or transcription factors), useful for treating cancer, e.g. colon cancer, gastric cancer, sarcoma, lymphoma or leukemia.

XX XX

PS Example 2; SEQ ID NO 2041; 134pp; English.

XX XX

CC The invention describes a purified polypeptide, which comprises a fragment of a kinase, phosphatase, protease, protease inhibitor, transporter, cytoskeletal protein, receptor or transcription factor. The polypeptide is useful as an immunogenic composition for eliciting in a mammal an immunogenic response directed against any of the purified polypeptide. The purified polypeptide, or the antibody that binds to this polypeptide, is useful for treating cancer. The polypeptide is also useful for identifying compounds that binds to a naturally processed class I or class II MHC-binding polypeptide. The polypeptides and polynucleotides are particularly useful for treating or preventing myeloma, colon cancer, gastric cancer, adenocarcinoma, sarcoma, melanoma, lymphoma or leukaemia. These are also useful for screening agents for treating the above mentioned diseases. This sequence represents an expressed protein tag (EPT) isolated from human tissue for translational profiling. Note: This sequence does not appear in the printed specification but was obtained in electronic format directly from WIPO at ftp.wipo.int/pub/published\_pct\_sequences

XX XX

Qy 1 LASLLGALAGPVLGLKECTRGSAVWCONVTASDCGAVKHCLQTVWKNKPTVKSLPCDIC 60

Db 7 LASLLGALAGPVLGLKECTRGSAVWCONVTASDCGAVKHCLQTVWKNKPTVKSLPCDIC 66

Qy 61 KDVVTAAGDMLKDNATTEELIYVLEKTCDWLPKPNMSASCKEIVDSYLPVLDIIRKEMS 120

Db 67 KDVVTAAGDMLKDNATTEELIYVLEKTCDWLPKPNMSASCKEIVDSYLPVLDIIRKEMS 126

Qy 121 RPGEVCSALNLCESLQKHLAEINHQKLESNKIPELDMEVAVPFMANIPILLYPQDQPR 180

Db 127 RPGEVCSALNLCESLQKHLAEINHQKLESNKIPELDMEVAVPFMANIPILLYPQDQPR 186

Qy 181 SKQPQKNGDVQCQICQWVTDIQTAVRTNSTFVQALVEHVKECDRLGPGWADICKNYS 240

Db 187 SKQPQKNGDVQCQICQWVTDIQTAVRTNSTFVQALVEHVKECDRLGPGWADICKNYS 246

Qy 241 QYSEIATQMMHMQPKEICALVGCDCEVKNPMQTLVPKAVSKNVIPALELVEPIKKHE 300

Db 247 QYSEIATQMMHMQPKEICALVGCDCEVKNPMQTLVPKAVSKNVIPALELVEPIKKHE 306

CC lymphoma or leukaemia. These are also useful for screening agents for  
CC treating the above mentioned diseases. This sequence represents an  
CC expressed protein tag (EPT) isolated from human tissue for translational  
CC profiling. Note: This sequence does not appear in the printed  
CC specification but was obtained in electronic format directly from WIPO at  
CC ftp.wipo.int/pub/published\_pct\_sequences  
XX  
SQ Sequence 524 AA;

Query Match 100.0%; Score 518; DB 6; Length 524;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 518; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 1 LASLLGAALAGPVILGLKECTRGSAVWQNVKTSADCGAVKHCLQTVWVKPTVKSILPCDIC 60  
Db 7 LASLLGAALAGPVILGLKECTRGSAVWQNVKTSADCGAVKHCLQTVWVKPTVKSILPCDIC 66  
Qy 61 KDVVTAAGDMLKONATEEELVLEKTCWMLPKPNMSASCKEIVDSYLPVILDIKIGEMS 120  
Db 67 KDVVTAAGDMLKONATEEELVLEKTCWMLPKPNMSASCKEIVDSYLPVILDIKIGEMS 126  
Qy 121 RPGEVCSALNLCESLQKHLAELNHQKLESNKIPELDMTEVAVPFWANIPLLLYPDGPR 180  
Db 127 RPGEVCSALNLCESLQKHLAELNHQKLESNKIPELDMTEVAVPFWANIPLLLYPDGPR 186  
Qy 181 SKPOKNGDVCODCIQWVTDIQTAVRTNSTFVQALVEHVKEECDRLGFGWADICKNYIS 240  
Db 187 SKPOKNGDVCODCIQWVTDIQTAVRTNSTFVQALVEHVKEECDRLGFGWADICKNYIS 246  
Qy 241 QYSEIAIQMMHMQPKIEICALVGFCDEVKEMPQTLVPAKASKNVIIPALBELVEPIKKE 300  
Db 247 QYSEIAIQMMHMQPKIEICALVGFCDEVKEMPQTLVPAKASKNVIIPALBELVEPIKKE 306  
Qy 301 VPAKSDVYCEVCEFLVKEVTKLIDNNKTEKILDAFDKMSCKLPKSLSEECQEVVDYGS 360  
Db 307 VPAKSDVYCEVCEFLVKEVTKLIDNNKTEKILDAFDKMSCKLPKSLSEECQEVVDYGS 366  
Qy 361 SILSILLEVSPELVCSMLHLCSTGRLPALTVHVTQPKDGGFCVCKKLVGLDRNLK 420  
Db 367 SILSILLEVSPELVCSMLHLCSTGRLPALTVHVTQPKDGGFCVCKKLVGLDRNLK 426  
Qy 421 STKQBIILAALCKGCSFLPDYQKQCDQFVAEYEPVLEILVEVMDPSFVCLKIGACPSAH 480  
Db 427 STKQBIILAALCKGCSFLPDYQKQCDQFVAEYEPVLEILVEVMDPSFVCLKIGACPSAH 486  
Qy 481 KPLLCTEKICWGPSYWCQNTETAACNAVEHCKRHVWN 518  
Db 487 KPLLCTEKICWGPSYWCQNTETAACNAVEHCKRHVWN 524

RESULT 10  
ABU05202  
ID ABU05202 standard; protein; 524 AA.  
XX AC ABU05202;  
XX DT 29-JAN-2003 (first entry)  
XX DE Human expressed protein tag (EPT) #1868.  
XX KW Translational profiling; expressed protein tag; EPT; kinase; phosphatase;  
KW protease; protease inhibitor; transporter; cytoskeletal protein;  
KW receptor; transcription factor; cancer; MHC;  
KW major histocompatibility complex; myeloma; colon cancer; gastric cancer;  
KW adenocarcinoma; sarcoma; melanoma; lymphoma; leukaemia.  
XX OS Homo sapiens.  
XX FN WO200278524-A2.  
XX PD 10-OCT-2002.  
XX PF 28-MAR-2002; 2002WO-US009671.

XX 28-MAR-2001; 2001US-0279495P.  
PR 21-MAY-2001; 2001US-0292544P.  
PR 08-AUG-2001; 2001US-0310801P.  
PR 01-OCT-2001; 2001US-0326370P.  
PR 04-DEC-2001; 2001US-0336780P.  
PR 20-FEB-2002; 2002US-0358985P.  
XX (ZYCO-) ZYCOS INC.  
XX Chicx RM, Tomlinson AJ, Urban RG;  
XX WPI; 2003-040607/03.  
XX New polypeptides (e.g. kinases, phosphatases, proteases, transporters,  
XX cytoskeletal proteins, receptors or transcription factors), useful for  
XX treating cancer, e.g. colon cancer, gastric cancer, sarcoma, lymphoma or  
XX leukemia.  
XX Example 2; SEQ ID NO 1868; 134pp; English.  
XX The invention describes a purified polypeptide, which comprises a  
XX fragment of a kinase, phosphatase, protease, or transcription inhibitor,  
XX transporter, cytoskeletal protein, receptor or transcription factor. The  
XX polypeptide is useful as an immunogenic composition for eliciting in a  
XX mammal an immunogenic response directed against any of the purified  
XX polypeptide. The purified polypeptide, or the antibody that binds to this  
XX polypeptide, is useful for treating cancer. The polypeptide is also  
XX useful for identifying compounds that binds to a naturally processed  
XX class I or class II MHC-binding polypeptide. The polypeptides and  
XX polynucleotides are particularly useful for treating or preventing  
XX myeloma, colon cancer, gastric cancer, adenocarcinoma, sarcoma, melanoma,  
XX lymphoma or leukaemia. These are also useful for screening agents for  
XX treating the above mentioned diseases. This sequence represents an  
XX expressed protein tag (EPT) isolated from human tissue for translational  
XX profiling. Note: This sequence does not appear in the printed  
XX specification but was obtained in electronic format directly from WIPO at  
XX ftp.wipo.int/pub/published\_pct\_sequences  
XX Sequence 524 AA;

Query Match 100.0%; Score 518; DB 6; Length 524;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 518; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 1 LASLLGAALAGPVILGLKECTRGSAVWQNVKTSADCGAVKHCLQTVWVKPTVKSILPCDIC 60  
Db 7 LASLLGAALAGPVILGLKECTRGSAVWQNVKTSADCGAVKHCLQTVWVKPTVKSILPCDIC 66  
Qy 61 KDVVTAAGDMLKONATEEELVLEKTCWMLPKPNMSASCKEIVDSYLPVILDIKIGEMS 120  
Db 67 KDVVTAAGDMLKONATEEELVLEKTCWMLPKPNMSASCKEIVDSYLPVILDIKIGEMS 126  
Qy 121 RPGEVCSALNLCESLQKHLAELNHQKLESNKIPELDMTEVAVPFWANIPLLLYPDGPR 180  
Db 127 RPGEVCSALNLCESLQKHLAELNHQKLESNKIPELDMTEVAVPFWANIPLLLYPDGPR 186  
Qy 181 SKPOKNGDVCODCIQWVTDIQTAVRTNSTFVQALVEHVKEECDRLGFGWADICKNYIS 240  
Db 187 SKPOKNGDVCODCIQWVTDIQTAVRTNSTFVQALVEHVKEECDRLGFGWADICKNYIS 246  
Qy 241 QYSEIAIQMMHMQPKIEICALVGFCDEVKEMPQTLVPAKASKNVIIPALBELVEPIKKE 300  
Db 247 QYSEIAIQMMHMQPKIEICALVGFCDEVKEMPQTLVPAKASKNVIIPALBELVEPIKKE 306  
Qy 301 VPAKSDVYCEVCEFLVKEVTKLIDNNKTEKILDAFDKMSCKLPKSLSEECQEVVDYGS 360  
Db 307 VPAKSDVYCEVCEFLVKEVTKLIDNNKTEKILDAFDKMSCKLPKSLSEECQEVVDYGS 366  
Qy 361 SILSILLEVSPELVCSMLHLCSTGRLPALTVHVTQPKDGGFCVCKKLVGLDRNLK 420  
Db 367 SILSILLEVSPELVCSMLHLCSTGRLPALTVHVTQPKDGGFCVCKKLVGLDRNLK 426

Qy 421 STKQEIILAALEKGCFLPDYQKQCDQFVAEYEPVLIEILVEVMDPSFVCLKIGACPSAH 480  
Db 427 STKQEIILAALEKGCFLPDYQKQCDQFVAEYEPVLIEILVEVMDPSFVCLKIGACPSAH 486  
Qy 481 KPLLGTKEKCIWGPSYWCNTETAQAQCNVAHECHKRHWVN 518  
Db 487 KPLLGTKEKCIWGPSYWCNTETAQAQCNVAHECHKRHWVN 524

RESULT 11  
ABU05214  
ID ABU05214 standard; protein; 524 AA.  
AC ABU05214;  
XX  
DT 29-JAN-2003 (first entry)  
XX  
DE Human expressed protein tag (EPT) #1880.  
XX  
KW Translational profiling; expressed protein tag; EPT; kinase; phosphatase;  
KW protease; protease inhibitor; transporter; cytoskeletal protein;  
KW receptor; transcription factor; cancer; MHC;  
KW major histocompatibility complex; myeloma; colon cancer; gastric cancer;  
KW adenocarcinoma; sarcoma; melanoma; lymphoma; leukaemia.  
XX  
OS Homo sapiens.  
XX  
XX WO200278524-A2.  
XX  
PD 10-OCT-2002.  
XX  
XX  
XX 28-MAR-2002; 2002WO-US009671.  
XX  
XX 28-MAR-2001; 2001US-0279495P.  
XX  
XX 21-MAY-2001; 2001US-0292544P.  
XX  
XX 08-AUG-2001; 2001US-0310801P.  
XX  
XX 01-OCT-2001; 2001US-0326370P.  
XX  
XX 04-DEC-2001; 2001US-0336780P.  
XX  
XX 20-FEB-2002; 2002US-0358985P.  
XX  
XX (ZYCO-) ZYCOS INC.  
XX  
XX Chicx RM, Tomlinson AJ, Urban RG;  
XX  
XX WPI; 2003-040607/03.  
XX  
XX New polypeptides (e.g. kinases, phosphatases, proteases, transporters,  
PT cytoskeletal proteins, receptors or transcription factors), useful for  
PT treating cancer, e.g. colon cancer, gastric cancer, sarcoma, lymphoma or  
PT leukemia.  
XX  
XX Example 2; SEQ ID NO 1880; 134pp; English.  
XX  
XX The invention describes a purified polypeptide, which comprises a  
CC fragment of a kinase, phosphatase, protease, protease inhibitor,  
CC transporter, cytoskeletal protein, receptor or transcription factor. The  
CC polypeptide is useful as an immunogenic composition for eliciting in a  
CC mammal an immunogenic response directed against any of the purified  
CC polypeptide. The purified polypeptide, or the antibody that binds to this  
CC polypeptide, is useful for treating cancer. The polypeptide is also  
CC useful for identifying compounds that binds to a naturally processed  
CC class I or class II MHC-binding polypeptide. The polypeptides and  
CC polynucleotides are particularly useful for treating or preventing  
CC myeloma, colon cancer, gastric cancer, adenocarcinoma, sarcoma, melanoma,  
CC lymphoma or leukaemia. These are also useful for screening agents for  
CC treating the above mentioned diseases. This sequence represents an  
CC expressed protein tag (EPT) isolated from human tissue for translational  
CC profiling. Note: This sequence does not appear in the printed  
CC specification but was obtained in electronic format directly from WIPO at  
XX ftp.wipo.int/pub/published\_pct\_sequences  
XX  
XX Sequence 524 AA;

Query Match 100.0%; Score 518; DB 6; Length 524;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 518; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 1 LASLIGAAALAGPVILGLKECTRGSAVWCVNKTASDCGAVKHCLQTVWKNKPTVKSPLCDIC 60  
Db 7 LASLIGAAALAGPVILGLKECTRGSAVWCVNKTASDCGAVKHCLQTVWKNKPTVKSPLCDIC 66  
Qy 61 KDVTTAAGDMLKDNATEEEILVYLEKTCMDLPPNMSASCKEIVDSYLPVLDIILKEMS 120  
Db 67 KDVTTAAGDMLKDNATEEEILVYLEKTCMDLPPNMSASCKEIVDSYLPVLDIILKEMS 126  
Qy 121 RPGEVCSALNLCESLQKHLAELNHQKLESNKIPELDMTEVVAFFMANIPLLYPQDGP 180  
Db 127 RPGEVCSALNLCESLQKHLAELNHQKLESNKIPELDMTEVVAFFMANIPLLYPQDGP 186  
Qy 181 SKPQKNDGVDCCDCIQMVTDIQTAVRTNSTFFQALVEHVKEBCDRGLGPGMADICKNYIS 240  
Db 187 SKPQKNDGVDCCDCIQMVTDIQTAVRTNSTFFQALVEHVKEBCDRGLGPGMADICKNYIS 246  
Qy 241 QYSEIAIQMMHMQPKEICALVGFCDVEKEMPQTLVPAKVASKNVIPALELVEPIKKHE 300  
Db 247 QYSEIAIQMMHMQPKEICALVGFCDVEKEMPQTLVPAKVASKNVIPALELVEPIKKHE 306  
Qy 301 VPAKSDVYCEVCFVLKVEVTKLIDNNKTEKEILDAFDMCSKLPKLSSEBCQEVVDYGS 360  
Db 307 VPAKSDVYCEVCFVLKVEVTKLIDNNKTEKEILDAFDMCSKLPKLSSEBCQEVVDYGS 366  
Qy 361 SILSILILEVSPVLVCSMLHLCSTGRLPALTVHVTQPKDGGFCEVCKLVGYLDRLNLEKN 420  
Db 367 SILSILILEVSPVLVCSMLHLCSTGRLPALTVHVTQPKDGGFCEVCKLVGYLDRLNLEKN 426  
Qy 421 STKQEIILAALEKGCFLPDYQKQCDQFVAEYEPVLIEILVEVMDPSFVCLKIGACPSAH 480  
Db 427 STKQEIILAALEKGCFLPDYQKQCDQFVAEYEPVLIEILVEVMDPSFVCLKIGACPSAH 486  
Qy 481 KPLLGTKEKCIWGPSYWCNTETAQAQCNVAHECHKRHWVN 518  
Db 487 KPLLGTKEKCIWGPSYWCNTETAQAQCNVAHECHKRHWVN 524

RESULT 12  
ABU05215  
ID ABU05215 standard; protein; 524 AA.  
XX  
AC ABU05215;  
XX  
DT 29-JAN-2003 (first entry)  
XX  
DE Human expressed protein tag (EPT) #1881.  
XX  
KW Translational profiling; expressed protein tag; EPT; kinase; phosphatase;  
KW protease; protease inhibitor; transporter; cytoskeletal protein;  
KW receptor; transcription factor; cancer; MHC;  
KW major histocompatibility complex; myeloma; colon cancer; gastric cancer;  
KW adenocarcinoma; sarcoma; melanoma; lymphoma; leukaemia.  
XX  
OS Homo sapiens.  
XX  
XX WO200278524-A2.  
XX  
XX 10-OCT-2002.  
XX  
XX 28-MAR-2002; 2002WO-US009671.  
XX  
XX 28-MAR-2001; 2001US-0279495P.  
XX  
XX 21-MAY-2001; 2001US-0292544P.  
XX  
XX 08-AUG-2001; 2001US-0310801P.  
XX  
XX 01-OCT-2001; 2001US-0326370P.  
XX  
XX 04-DEC-2001; 2001US-0336780P.  
XX  
XX 20-FEB-2002; 2002US-0358985P.  
XX  
XX (ZYCO-) ZYCOS INC.  
XX  
XX



XX Chic RM, Tomlinson AJ, Urban RG;  
XX WPI; 2003-040607/03.  
XX  
XX New polypeptides (e.g. kinases, phosphatases, proteases, transporters,  
XX cytoskeletal proteins, receptors or transcription factors), useful for  
XX treating cancer, e.g. colon cancer, gastric cancer, sarcoma, lymphoma or  
XX leukemia.  
XX  
XX Example 2; SEQ ID NO 1881; 134pp; English.  
XX  
XX The invention describes a purified polypeptide, which comprises a  
XX fragment of a kinase, phosphatase, protease, protease inhibitor,  
XX transporter, cytoskeletal protein, receptor or transcription factor. The  
XX polypeptide is useful as an immunogenic composition for eliciting in a  
XX mammal an immunogenic response directed against any of the purified  
XX polypeptide. The purified polypeptide, or the antibody that binds to this  
XX polypeptide, is useful for treating cancer. The polypeptide is also  
XX useful for identifying compounds that binds to a naturally processed  
XX class I or class II MHC-binding polypeptide. The polypeptides and  
XX polynucleotides are particularly useful for treating or preventing  
XX myeloma, colon cancer, gastric cancer, adenocarcinoma, melanoma,  
XX lymphoma or leukaemia. These are also useful for screening agents for  
XX treating the above mentioned diseases. This sequence represents an  
XX expressed protein tag (EPT) isolated from human tissue for translational  
XX profiling. Note: This sequence does not appear in the printed  
XX specification but was obtained in electronic format directly from WIPO at  
XX ftp.wipo.int/pub/published\_pct\_sequences  
XX Sequence 524 AA;

Query Match 100.0%; Score 518; DB 6; Length 524;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 518; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 LASLLGALAGPVLGLKECTRGSAVQCNVKTASDCGAVKHCLQTVNKNPTVKSLPCDIC 60  
Db 7 LASLLGALAGPVLGLKECTRGSAVQCNVKTASDCGAVKHCLQTVNKNPTVKSLPCDIC 66  
QY 61 KDVTAAAGDMLKDNATTEEILVLEKTCDWLPKPNMSASCKEIVDSYLPVLDIIGEMS 120  
Db 67 KDVTAAAGDMLKDNATTEEILVLEKTCDWLPKPNMSASCKEIVDSYLPVLDIIGEMS 126  
QY 121 RPEVCSALNLCESLOKHLAELNHQKLESNKIPELDMEVAPFMANIPLLLYPDQGR 180  
Db 127 RPEVCSALNLCESLQKHLAELNHQKLESNKIPELDMEVAPFMANIPLLLYPDQGR 186  
QY 181 SKPQPKNGDVQCQICQWVTDIQTAVRTNSTFVQALVEHVKEECDRLGPGMADICKNYIS 240  
Db 187 SKPQPKNGDVQCQICQWVTDIQTAVRTNSTFVQALVEHVKEECDRLGPGMADICKNYIS 246  
QY 241 QYSEIAIQMMHMQPKIEICALVGFCDVEKMPQTLVPAKVASKNIPALVELVEPIKKEH 300  
Db 247 QYSEIAIQMMHMQPKIEICALVGFCDVEKMPQTLVPAKVASKNIPALVELVEPIKKEH 306  
QY 301 VPAKSDVYCEVCEFLVKEVTKLIDNNKTEKIELDAFDKMSKLPKSLSECCQEVDTYGS 360  
Db 307 VPAKSDVYCEVCEFLVKEVTKLIDNNKTEKIELDAFDKMSKLPKSLSECCQEVDTYGS 366  
QY 361 SILSILLEEVSPELVCSMLHCLSGTRLPALTVHTVQPKDGGFCEVCCKLVGYLDRLNLEKN 420  
Db 367 SILSILLEEVSPELVCSMLHCLSGTRLPALTVHTVQPKDGGFCEVCCKLVGYLDRLNLEKN 426  
QY 421 STKOEIILAALKGCSFLPDYQKQCOQFVAEYEPVLIEILVEVMDPSFVCLKIGACPSAH 480  
Db 427 STKOEIILAALKGCSFLPDYQKQCOQFVAEYEPVLIEILVEVMDPSFVCLKIGACPSAH 486  
QY 481 KPLLGTKECIWGPSYMCQNTETAACQNAVEHCKRHVWN 518  
Db 487 KPLLGTKECIWGPSYMCQNTETAACQNAVEHCKRHVWN 524

RESULT 13  
ABU05199  
ID ABU05199 standard; protein; 524 AA.  
XX  
XX AC ABU05199;  
XX  
XX 29-JAN-2003 (first entry)  
DT  
XX Human expressed protein tag (EPT) #1865.  
DE  
XX  
XX Translational profiling; expressed protein tag; EPT; kinase; phosphatase;  
XX protease; protease inhibitor; transporter; cytoskeletal protein;  
XX receptor; transcription factor; cancer; MHC;  
XX major histocompatibility complex; myeloma; colon cancer; gastric cancer;  
XX adenocarcinoma; sarcoma; melanoma; lymphoma; leukaemia.  
XX  
XX Homo sapiens.  
OS  
XX  
XX W0200278524-A2.  
PN  
XX  
XX 10-OCT-2002.  
PD  
XX  
XX 28-MAR-2002; 2002WO-US009671.  
PF  
XX  
XX 28-MAR-2001; 2001US-0279495P.  
PR  
XX  
XX 21-MAY-2001; 2001US-0292544P.  
PR  
XX  
XX 08-AUG-2001; 2001US-0310801P.  
PR  
XX  
XX 01-OCT-2001; 2001US-0326370P.  
PR  
XX  
XX 04-DEC-2001; 2001US-0336780P.  
PR  
XX  
XX 20-FEB-2002; 2002US-0358985P.  
PR  
XX  
XX (ZYCO-) ZYCOS INC.  
PA  
XX  
XX Chic RM, Tomlinson AJ, Urban RG;  
PI  
XX  
XX WPI; 2003-040607/03.  
DR  
XX  
XX New polypeptides (e.g. kinases, phosphatases, proteases, transporters,  
XX cytoskeletal proteins, receptors or transcription factors), useful for  
XX treating cancer, e.g. colon cancer, gastric cancer, sarcoma, lymphoma or  
XX leukemia.  
XX  
XX Example 2; SEQ ID NO 1865; 134pp; English.  
XX  
XX The invention describes a purified polypeptide, which comprises a  
XX fragment of a kinase, phosphatase, protease, protease inhibitor,  
XX transporter, cytoskeletal protein, receptor or transcription factor. The  
XX polypeptide is useful as an immunogenic composition for eliciting in a  
XX mammal an immunogenic response directed against any of the purified  
XX polypeptide. The purified polypeptide, or the antibody that binds to this  
XX polypeptide, is useful for treating cancer. The polypeptide is also  
XX useful for identifying compounds that binds to a naturally processed  
XX class I or class II MHC-binding polypeptide. The polypeptides and  
XX polynucleotides are particularly useful for treating or preventing  
XX myeloma, colon cancer, gastric cancer, adenocarcinoma, sarcoma, melanoma,  
XX lymphoma or leukaemia. These are also useful for screening agents for  
XX treating the above mentioned diseases. This sequence represents an  
XX expressed protein tag (EPT) isolated from human tissue for translational  
XX profiling. Note: This sequence does not appear in the printed  
XX specification but was obtained in electronic format directly from WIPO at  
XX ftp.wipo.int/pub/published\_pct\_sequences  
XX Sequence 524 AA;

Query Match 100.0%; Score 518; DB 6; Length 524;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 518; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 LASLLGALAGPVLGLKECTRGSAVQCNVKTASDCGAVKHCLQTVNKNPTVKSLPCDIC 60  
Db 7 LASLLGALAGPVLGLKECTRGSAVQCNVKTASDCGAVKHCLQTVNKNPTVKSLPCDIC 66  
QY 61 KDVTAAAGDMLKDNATTEEILVLEKTCDWLPKPNMSASCKEIVDSYLPVLDIIGEMS 120



Db 67 KDVVTAAGDMLKDNATEEEILVYLEKTCDWLPKPNMSASCKEIVDSYLPVLDIIKEMS 126  
Qy 121 RPEVCSALNLCESLQKHLAEHLNKHQKLESNKIPELDMEVAVPFFMANIPLLLYPQDGP 180  
Db 127 RPEVCSALNLCESLQKHLAEHLNKHQKLESNKIPELDMEVAVPFFMANIPLLLYPQDGP 186  
Qy 181 SKPOKNDGVQCQICIQWVTDIQTAVRTNSTFVQALVEHVKEECDRGLGPGMADICKNYIS 240  
Db 187 SKPOKNDGVQCQICIQWVTDIQTAVRTNSTFVQALVEHVKEECDRGLGPGMADICKNYIS 246  
Qy 241 QYSEIAIQMMHMOPKEICALVGFCDVEKEMPMQTLVPAKVASKNVIPAELVEPIKKHE 300  
Db 247 QYSEIAIQMMHMOPKEICALVGFCDVEKEMPMQTLVPAKVASKNVIPAELVEPIKKHE 306  
Qy 301 VPAKSDVYCEVCEFLVKEVTKLIDNNKTEKILDAFDMKCSKLPKSLSEECQEVVDYGS 360  
Db 307 VPAKSDVYCEVCEFLVKEVTKLIDNNKTEKILDAFDMKCSKLPKSLSEECQEVVDYGS 366  
Qy 361 SILSILLEEVSPELVCSMLHLCSTGTRLPALTAVHTVQPKDGGFCVCKKLVGYLDRLNLEKN 420  
Db 367 SILSILLEEVSPELVCSMLHLCSTGTRLPALTAVHTVQPKDGGFCVCKKLVGYLDRLNLEKN 426  
Qy 421 STKQEIILAALEKGCFLPDYQKQCDQFVAEYEPVLIIEILVEVMDPSFVCLKIGACPSAH 480  
Db 427 STKQEIILAALEKGCFLPDYQKQCDQFVAEYEPVLIIEILVEVMDPSFVCLKIGACPSAH 486  
Qy 481 KPLLGTCKCIWGPSYWCQNTETAACQNAVEHCKRHHVN 518  
Db 487 KPLLGTCKCIWGPSYWCQNTETAACQNAVEHCKRHHVN 524  
RESULT 14  
ABU05212  
ID ABU05212 standard; protein; 524 AA.  
XX  
AC ABU05212;  
XX  
DT 29-JAN-2003 (first entry)  
XX  
DE Human expressed protein tag (EPT) #1878.  
KW Translational profiling; expressed protein tag; EPT; kinase; phosphatase;  
KW protease; protease inhibitor; transporter; cytoskeletal protein;  
KW receptor; transcription factor; cancer; MHC;  
KW major histocompatibility complex; myeloma; colon cancer; gastric cancer;  
KW adenocarcinoma; sarcoma; melanoma; lymphoma; leukaemia.  
XX Homo sapiens.  
XX  
XX WO200278524-A2.  
XX  
XX 10-OCT-2002.  
XX  
XX 28-MAR-2002; 2002WO-US009671.  
XX  
XX 28-MAR-2001; 2001US-0279495P.  
XX 21-MAY-2001; 2001US-0292544P.  
XX 08-AUG-2001; 2001US-0310801P.  
XX 01-OCT-2001; 2001US-0326370P.  
XX 04-DEC-2001; 2001US-0336780P.  
XX 20-FEB-2002; 2002US-0358985P.  
XX  
XX (ZYCO-) ZYCOs INC.  
XX  
XX Chicx RM, Tomlinson AJ, Urban RG;  
XX  
XX WPI; 2003-040607/03.  
XX  
XX New polypeptides (e.g. kinases, phosphatases, proteases, transporters,  
XX cytoskeletal proteins, receptors or transcription factors), useful for  
XX treating cancer, e.g. colon cancer, gastric cancer, sarcoma, lymphoma or  
XX leukemia.

XX Example 2; SEQ ID NO 1878; 134pp; English.  
XX  
CC The invention describes a purified polypeptide, which comprises a  
CC fragment of a kinase, phosphatase, protease, or transcription inhibitor,  
CC transporter, cytoskeletal protein, receptor or transcription factor. The  
CC polypeptide is useful as an immunogenic composition for eliciting in a  
CC mammal an immunogenic response directed against any of the purified  
CC polypeptide, the purified polypeptide, or the antibody that binds to this  
CC polypeptide, is useful for treating cancer. The polypeptide is also  
CC useful for identifying compounds that binds to a naturally processed  
CC class I or class II MHC-binding polypeptide. The polypeptides and  
CC polynucleotides are particularly useful for treating or preventing  
CC myeloma, colon cancer, gastric cancer, adenocarcinoma, sarcoma, melanoma,  
CC lymphoma or leukaemia. These are also useful for screening agents for  
CC treating the above mentioned diseases. This sequence represents an  
CC expressed protein tag (EPT) isolated from human tissue for translational  
CC profiling. Note: This sequence does not appear in the printed  
CC specification but was obtained in electronic format directly from WIPO at  
CC ftp.wipo.int/pub/published\_pct\_sequences  
XX  
SQ Sequence 524 AA;  
Query Match 100.0%; Score 518; DB 6; Length 524;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 518; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 1 LASLGAALAGPVILGLKECTRGSAVWCQNVKTSADCGAVKHCLQTVWKNKPTVKSLPCDIC 60  
Db 7 LASLGAALAGPVILGLKECTRGSAVWCQNVKTSADCGAVKHCLQTVWKNKPTVKSLPCDIC 66  
Qy 61 KDVVTAAGDMLKDNATEEEILVYLEKTCDWLPKPNMSASCKEIVDSYLPVLDIIKEMS 120  
Db 67 KDVVTAAGDMLKDNATEEEILVYLEKTCDWLPKPNMSASCKEIVDSYLPVLDIIKEMS 126  
Qy 121 RPEVCSALNLCESLQKHLAEHLNKHQKLESNKIPELDMEVAVPFFMANIPLLLYPQDGP 180  
Db 127 RPEVCSALNLCESLQKHLAEHLNKHQKLESNKIPELDMEVAVPFFMANIPLLLYPQDGP 186  
Qy 181 SKPOKNDGVQCQICIQWVTDIQTAVRTNSTFVQALVEHVKEECDRGLGPGMADICKNYIS 240  
Db 187 SKPOKNDGVQCQICIQWVTDIQTAVRTNSTFVQALVEHVKEECDRGLGPGMADICKNYIS 246  
Qy 241 QYSEIAIQMMHMOPKEICALVGFCDVEKEMPMQTLVPAKVASKNVIPAELVEPIKKHE 300  
Db 247 QYSEIAIQMMHMOPKEICALVGFCDVEKEMPMQTLVPAKVASKNVIPAELVEPIKKHE 306  
Qy 301 VPAKSDVYCEVCEFLVKEVTKLIDNNKTEKILDAFDMKCSKLPKSLSEECQEVVDYGS 360  
Db 307 VPAKSDVYCEVCEFLVKEVTKLIDNNKTEKILDAFDMKCSKLPKSLSEECQEVVDYGS 366  
Qy 361 SILSILLEEVSPELVCSMLHLCSTGTRLPALTAVHTVQPKDGGFCVCKKLVGYLDRLNLEKN 420  
Db 367 SILSILLEEVSPELVCSMLHLCSTGTRLPALTAVHTVQPKDGGFCVCKKLVGYLDRLNLEKN 426  
Qy 421 STKQEIILAALEKGCFLPDYQKQCDQFVAEYEPVLIIEILVEVMDPSFVCLKIGACPSAH 480  
Db 427 STKQEIILAALEKGCFLPDYQKQCDQFVAEYEPVLIIEILVEVMDPSFVCLKIGACPSAH 486  
Qy 481 KPLLGTCKCIWGPSYWCQNTETAACQNAVEHCKRHHVN 518  
Db 487 KPLLGTCKCIWGPSYWCQNTETAACQNAVEHCKRHHVN 524  
RESULT 15  
ABU05213  
ID ABU05213 standard; protein; 524 AA.  
XX  
XX AC ABU05213;  
XX  
XX 29-JAN-2003 (first entry)  
XX  
XX Human expressed protein tag (EPT) #1879.  
DE

Translational profiling; expressed protein tag; EPT; kinase; phosphatase; protease; protease inhibitor; transporter; cytoskeletal protein; receptor; transcription factor; cancer; MHC; major histocompatibility complex; myeloma; colon cancer; gastric cancer; adenocarcinoma; sarcoma; melanoma; lymphoma; leukaemia.

Homo sapiens.

WO200278524-A2.

10-OCT-2002.

28-MAR-2002; 2002WO-US009671.

28-MAR-2001; 2001US-0279495P.

21-MAY-2001; 2001US-0292544P.

08-AUG-2001; 2001US-0310801P.

01-OCT-2001; 2001US-0326370P.

04-DEC-2001; 2001US-0336780P.

20-FEB-2002; 2002US-0359895P.

(ZYCO-) ZYCOS INC.

Chicz RM, Tomlinson AJ, Urban RG; WPI; 2003-040607/03.

New polypeptides (e.g. kinases, phosphatases, proteases, transporters, cytoskeletal proteins, receptors or transcription factors), useful for treating cancer, e.g. colon cancer, gastric cancer, sarcoma, lymphoma or leukemia.

Example 2; SEQ ID NO 1879; 134pp; English.

The invention describes a purified polypeptide, which comprises a fragment of a kinase, phosphatase, protease, protease inhibitor, transporter, cytoskeletal protein, receptor or transcription factor. The polypeptide is useful as an immunogenic composition for eliciting in a mammal an immunogenic response directed against any of the purified polypeptide. The purified polypeptide, or the antibody that binds to this polypeptide, is useful for treating cancer. The polypeptide is also useful for identifying compounds that binds to a naturally processed class I or class II MHC-binding polypeptide. The polypeptides and polynucleotides are particularly useful for treating or preventing myeloma, colon cancer, gastric cancer, adenocarcinoma, sarcoma, melanoma, lymphoma or leukaemia. These are also useful for screening agents for treating the above mentioned diseases. This sequence represents an expressed protein tag (EPT) isolated from human tissue for translational profiling. Note: This sequence does not appear in the printed specification but was obtained in electronic format directly from WIPO at ftp.wipo.int/pub/published\_pct\_sequences

Sequence 524 AA;

Query Match 100.0%; Score 518; DB 6; Length 524;  
 Best Local Similarity 100.0%; Pred. No. 0;  
 Matches 518; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LASLIGALAGFVLGLKECTRGSAVWCQNVKTA DCGAVKHCLQTVWNKPTVKSLPCDIC 60  
 Db 7 LASLIGALAGFVLGLKECTRGSAVWCQNVKTA DCGAVKHCLQTVWNKPTVKSLPCDIC 66

QY 61 KDVTVAQDMLKDNATEEEILVLEKTCDWLPKPNMSASCKEIVDSYLPVILDIKEMS 120  
 Db 67 KDVTVAQDMLKDNATEEEILVLEKTCDWLPKPNMSASCKEIVDSYLPVILDIKEMS 156

QY 121 RFGVCSALNLCESLQKHAELNHQKLESNKIPELDMTEVAVPMANIPILLYPDGGR 180  
 Db 127 RFGVCSALNLCESLQKHAELNHQKLESNKIPELDMTEVAVPMANIPILLYPDGGR 186

QY 191 SKPQPKNGVCDQCIQMTDITQTVRTNSTVQALVEHVKEECDRLGPGMADICKNYIS 240

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 5, 2004, 13:28:28 ; Search time 15.405 Seconds  
(without alignments)  
1630.129 Million cell updates/sec

Title: US-09-743-684A-1\_COPY\_7\_524

Perfect score: 518  
Sequence: 1 LASLLGALAGPVLGLKCT.....NTETAQCNAVEHCKRHWN 518

Scoring table: OLIGO  
Gapop 60.0 , Gapext 60.0

Searched: 389414 seqs, 51625971 residues

Word size : 0

Total number of hits satisfying chosen parameters: 389414

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : Issued Patents AA.\*  
1: /cgn2\_6/ptodata/2/iaa/5A\_COMB.pep.\*  
2: /cgn2\_6/ptodata/2/iaa/5B\_COMB.pep.\*  
3: /cgn2\_6/ptodata/2/iaa/6A\_COMB.pep.\*  
4: /cgn2\_6/ptodata/2/iaa/6B\_COMB.pep.\*  
5: /cgn2\_6/ptodata/2/iaa/PCTUS\_COMB.pep.\*  
6: /cgn2\_6/ptodata/2/iaa/backfiles.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	518	100.0	524	4	US-09-352-548-1
2	224	43.2	523	1	US-08-100-247-2
3	224	43.2	523	1	US-08-232-513A-3
4	224	43.2	523	4	US-08-756-031-2
5	187	36.1	523	1	US-08-483-146A-2
6	187	36.1	523	1	US-08-484-594A-2
7	187	36.1	523	4	US-08-076-258A-2
8	81	15.6	81	4	US-09-352-548-2
9	80	15.4	80	1	US-08-100-247-3
10	80	15.4	80	1	US-08-483-146A-3
11	80	15.4	80	1	US-08-232-513A-4
12	80	15.4	80	1	US-08-484-594A-3
13	80	15.4	80	4	US-08-076-258A-3
14	80	15.4	80	4	US-08-756-031-3
15	75	14.5	80	2	US-08-584-671-15
16	75	14.5	80	3	US-09-027-376-15
17	69	13.3	80	3	US-09-094-192-15
18	61	11.8	61	2	US-08-584-671-13
19	61	11.8	61	3	US-09-027-376-13
20	61	11.8	61	3	US-09-094-192-13
21	61	11.8	69	3	US-09-268-070-2
22	22	4.2	22	1	US-08-100-247-1
23	22	4.2	22	1	US-08-483-146A-1
24	22	4.2	22	1	US-08-483-146A-11
25	22	4.2	22	1	US-08-232-513A-1
26	22	4.2	22	1	US-08-232-513A-20
27	22	4.2	22	1	US-08-484-594A-1

28	22	4.2	22	1	US-08-484-594A-11	Sequence 11, Appl
29	22	4.2	22	3	US-09-231-159-1	Sequence 1, Appl
30	22	4.2	22	3	US-08-611-307-1	Sequence 1, Appl
31	22	4.2	22	4	US-09-148-030-1	Sequence 1, Appl
32	22	4.2	22	4	US-09-076-258A-1	Sequence 1, Appl
33	22	4.2	22	4	US-09-076-258A-11	Sequence 11, Appl
34	22	4.2	22	4	US-08-756-031-1	Sequence 1, Appl
35	18	3.5	18	4	US-09-352-548-17	Sequence 17, Appl
36	17	3.3	17	4	US-09-352-548-14	Sequence 14, Appl
37	17	3.3	17	4	US-09-352-548-15	Sequence 15, Appl
38	17	3.3	18	1	US-08-100-247-5	Sequence 5, Appl
39	17	3.3	18	1	US-08-483-146A-5	Sequence 5, Appl
40	17	3.3	18	1	US-08-232-513A-6	Sequence 6, Appl
41	17	3.3	18	1	US-08-484-594A-5	Sequence 5, Appl
42	17	3.3	18	3	US-09-231-159-20	Sequence 20, Appl
43	17	3.3	18	3	US-08-611-307-20	Sequence 20, Appl
44	17	3.3	18	4	US-09-148-030-2	Sequence 2, Appl
45	17	3.3	18	4	US-09-076-258A-5	Sequence 5, Appl

ALIGNMENTS

RESULT 1  
US-09-352-548-1  
; Sequence 1, Application US/09352548  
; Patent No. 6500431  
; GENERAL INFORMATION:  
; APPLICANT: Gili, Parkash S.  
; APPLICANT: Parkash S. Gili, M.D., Inc.  
; TITLE OF INVENTION: No. 6500431ei Inhibitors of Angiogenesis and Tumor Growth  
; FILE REFERENCE: 017986-000410US  
; CURRENT APPLICATION NUMBER: US/09/352,548  
; CURRENT FILING DATE: 1999-07-12  
; EARLIER APPLICATION NUMBER: US 60/092,647  
; EARLIER FILING DATE: 1998-07-13  
; NUMBER OF SEQ ID NOS: 59  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 1  
; LENGTH: 524  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
; FEATURE:  
; OTHER INFORMATION: prosaposin  
; FEATURE:  
; NAME/KEY: PEPTIDE  
; LOCATION: (195)..(275)  
; OTHER INFORMATION: Saposin B  
US-09-352-548-1

Query Match	100.0%;	Score 518;	DB 4;	Length 524;
Best Local Similarity	100.0%;	Pred. No. 0;		
Matches 518;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
Qy	1	LASLLGALAGPVLGLKCTRGSAVQCQNVKTSADCGAVKHCLQTVNWKPTVKSILPCDIC	60	
Db	7	LASLLGALAGPVLGLKCTRGSAVQCQNVKTSADCGAVKHCLQTVNWKPTVKSILPCDIC	66	
Qy	61	KDVVTAAGDMLKDNATEEILVYLEKCDMLPKNMSASCKEIVDSYLPVILDIKIGEMS	120	
Db	67	KDVVTAAGDMLKDNATEEILVYLEKCDMLPKNMSASCKEIVDSYLPVILDIKIGEMS	126	
Qy	121	RPGEVCSALNLCESLQKHAEIHNHOKLESNKIPELDMTEVVAFPMANIPLLLYPDQDGR	180	
Db	127	RPGEVCSALNLCESLQKHAEIHNHOKLESNKIPELDMTEVVAFPMANIPLLLYPDQDGR	186	
Qy	181	SKPQKDNQGVQCDCIQMVTDIQTAVRTNSTFVQALVEHVKEECDRLPGMGADICKNYIS	240	
Db	187	SKPQKDNQGVQCDCIQMVTDIQTAVRTNSTFVQALVEHVKEECDRLPGMGADICKNYIS	246	
Qy	241	QYSEIAIQMMHMQPKEICALVGFCDEKEMPMQTLVPKVASKNVPALELVEPIKKHE	300	
Db	247	QYSEIAIQMMHMQPKEICALVGFCDEKEMPMQTLVPKVASKNVPALELVEPIKKHE	306	

QY 301 VPAKSDVYCEVCEFLVKEVTKLIDNNKTEKEILDAFDKMKSKLPKSLSECEQVVDITYGS 360  
DB 307 VPAKSDVYCEVCEFLVKEVTKLIDNNKTEKEILDAFDKMKSKLPKSLSECEQVVDITYGS 366  
QY 361 SILSILILEEVSPPELVCSMLHLCSTGTRLPALTVHTVTPKDGCGFCEVCKLVGLVDRNLEKN 420  
DB 367 SILSILILEEVSPPELVCSMLHLCSTGTRLPALTVHTVTPKDGCGFCEVCKLVGLVDRNLEKN 426  
QY 421 STKQEIILAALEKSGFLPDYQKQCDQFVAEYEPVLIIEILVEVMDPSFVCLKIGACPSAH 480  
DB 427 STKQEIILAALEKSGFLPDYQKQCDQFVAEYEPVLIIEILVEVMDPSFVCLKIGACPSAH 486  
QY 481 KPLLGTCKIWPSPYWCNTETAACQNAVEHCKRHVWN 518  
DB 487 KPLLGTCKIWPSPYWCNTETAACQNAVEHCKRHVWN 524

## RESULT 2

US-08-100-247-2  
; Sequence 2, Application US/08100247  
; Patent No. 5571787  
; GENERAL INFORMATION:  
; APPLICANT: O'BRIEN, JOHN S.  
; APPLICANT: KISHIMOTO, YASUO  
; TITLE OF INVENTION: PROSAPOSIN AS A NEUROTROPHIC FACTOR  
; NUMBER OF SEQUENCES: 5  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: KNOBBE, MARTENS, OLSON AND BEAR  
; STREET: 620 NEWPORT CENTER DRIVE SIXTEENTH FLOOR  
; CITY: NEWPORT BEACH  
; STATE: CA  
; COUNTRY: USA  
; ZIP: 92660

COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent in Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/100.247  
; FILING DATE: 19930730  
; CLASSIFICATION: 514

ATTORNEY/AGENT INFORMATION:  
; NAME: Israelsen, Ned A.  
; REGISTRATION NUMBER: 29,655  
; REFERENCE/DOCKET NUMBER: O'Brien.002A  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 619-235-8550  
; TELEFAX: 619-235-0176  
; INFORMATION FOR SEQ ID NO: 2:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 523 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
; HYPOTHEICAL: NO  
; ANTI-SENSE: NO  
; FRAGMENT TYPE: N-terminal  
; IMMEDIATE SOURCE:  
; CLONE: PROSAPOSIN

US-08-100-247-2

Query Match 43.2%; Score 224; DB 1; Length 523;  
Best Local Similarity 100.0%; Pred. No. 1e-207;  
Matches 224; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 295 PIKKEVPAKSDVYCEVCEFLVKEVTKLIDNNKTEKEILDAFDKMKSKLPKSLSECEQV 354  
DB 300 PIKKEVPAKSDVYCEVCEFLVKEVTKLIDNNKTEKEILDAFDKMKSKLPKSLSECEQV 359  
QY 355 VDTYGSILSILILEEVSPPELVCSMLHLCSTGTRLPALTVHTVTPKDGCGFCEVCKLVGLYLD 414

DB 360 VDTYGSILSILILEEVSPPELVCSMLHLCSTGTRLPALTVHTVTPKDGCGFCEVCKLVGLYLD 419  
QY 415 RNLEKNTKQEIILAALEKSGFLPDYQKQCDQFVAEYEPVLIIEILVEVMDPSFVCLKIG 474  
DB 420 RNLEKNTKQEIILAALEKSGFLPDYQKQCDQFVAEYEPVLIIEILVEVMDPSFVCLKIG 479  
QY 475 ACPSAHKPLLGTCKIWPSPYWCNTETAACQNAVEHCKRHVWN 518  
DB 480 ACPSAHKPLLGTCKIWPSPYWCNTETAACQNAVEHCKRHVWN 523

## RESULT 3

US-08-232-513A-3  
; Sequence 3, Application US/08232513A  
; Patent No. 5700909  
; GENERAL INFORMATION:  
; APPLICANT: O'Brien, John S.  
; TITLE OF INVENTION: ProsaPOSIN and Cytokine-Derived Peptides  
; TITLE OF INVENTION: as Therapeutic Agents  
; NUMBER OF SEQUENCES: 20  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Cambell & Flores LLP  
; STREET: 4370 La Jolla Village Drive, Suite 700  
; CITY: San Diego  
; STATE: California  
; COUNTRY: USA  
; ZIP: 92122

COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent in Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/232.513A  
; FILING DATE: 21-APR-1994  
; CLASSIFICATION: 514

PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/100.247  
; FILING DATE: 30-JUL-1993  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Cambell, Cathryn A.  
; REGISTRATION NUMBER: 31,815  
; REFERENCE/DOCKET NUMBER: P-UD 1643  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (619) 535-9001  
; TELEFAX: (619) 535-8949  
; INFORMATION FOR SEQ ID NO: 3:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 523 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; FEATURE:  
; NAME/KEY: Protein  
; LOCATION: 1..523  
; OTHER INFORMATION: /label= Hum\_prosaPOSIN  
US-08-232-513A-3

Query Match 43.2%; Score 224; DB 1; Length 523;  
Best Local Similarity 100.0%; Pred. No. 1e-207;  
Matches 224; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 295 PIKKEVPAKSDVYCEVCEFLVKEVTKLIDNNKTEKEILDAFDKMKSKLPKSLSECEQV 354  
DB 300 PIKKEVPAKSDVYCEVCEFLVKEVTKLIDNNKTEKEILDAFDKMKSKLPKSLSECEQV 359  
QY 355 VDTYGSILSILILEEVSPPELVCSMLHLCSTGTRLPALTVHTVTPKDGCGFCEVCKLVGLYLD 414  
DB 360 VDTYGSILSILILEEVSPPELVCSMLHLCSTGTRLPALTVHTVTPKDGCGFCEVCKLVGLYLD 419  
QY 415 RNLEKNTKQEIILAALEKSGFLPDYQKQCDQFVAEYEPVLIIEILVEVMDPSFVCLKIG 474

Db 420 RNLEKNTKQBIILAALAKGCSFLPDPYQKQCDQFVAEYEPVLIELVEVMDPSFVCLKIG 479

Qy 475 ACPSAHKPLLGTEKCIWGPSYWCNTETATAACNAVEHCKRHVWN 518  
|||||

Db 480 ACPSAHKPLLGTEKCIWGPSYWCNTETATAACNAVEHCKRHVWN 523  
|||||

RESULT 4

US-08-756-031-2

; Sequence 2, Application US/08756031

; Patent No. 6590074

; GENERAL INFORMATION:

; APPLICANT: O'BRIEN, JOHN S.

; APPLICANT: KISHIMOTO, YASUO

; TITLE OF INVENTION: PROSAPOSIN AS A NEUROTROPHIC FACTOR

; NUMBER OF SEQUENCES: 5

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: KNOBBE, MARTENS, OLSON AND BEAR

; STREET: 620 NEWPORT CENTER DRIVE SIXTEENTH FLOOR

; CITY: NEWPORT BEACH

; STATE: CA

; COUNTRY: USA

; ZIP: 92660

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: Patent In Release #1.0, Version #1.25

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/756,031

; FILING DATE: 26-NOV-1996

; CLASSIFICATION: 514

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: 08/100,247

; FILING DATE: 30-JUL-1993

; ATTORNEY/AGENT INFORMATION:

; NAME: Israelsen, Ned A.

; REGISTRATION NUMBER: 29,655

; REFERENCE/DOCKET NUMBER: O'BRIEN.002A

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: 619-235-8550

; TELEFAX: 619-235-0176

; INFORMATION FOR SEQ ID NO: 2:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 523 amino acids

; TYPE: amino acid

; STRANDEDNESS: single

; TOPOLOGY: linear

; MOLECULE TYPE: peptide

; HYPOTHETICAL: NO

; ANTI-SENSE: NO

; FRAGMENT TYPE: N-terminal

; IMMEDIATE SOURCE:

; CLONE: PROSAPOSIN

US-08-756-031-2

Query Match 43.2%; Score 224; DB 4; Length 523;

Best Local Similarity 100.0%; Pred. No. 1e-207;

Matches 224; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 295 PIKKEHPAKSDVYCEVCEFLVKEVTKLIDNNKTEKILDAFDKVCCKLPKSLSEECQEV 354  
|||||

Db 300 PIKKEHPAKSDVYCEVCEFLVKEVTKLIDNNKTEKILDAFDKVCCKLPKSLSEECQEV 359  
|||||

Qy 355 VDTYGSSTLSILLBEVSPELVCSMLHLCSTGTRLPALTTHVTQPKDGGCFCECKLVGYLD 414  
|||||

Db 360 VDTYGSSTLSILLBEVSPELVCSMLHLCSTGTRLPALTTHVTQPKDGGCFCECKLVGYLD 419  
|||||

Qy 415 RNLEKNTKQBIILAALAKGCSFLPDPYQKQCDQFVAEYEPVLIELVEVMDPSFVCLKIG 474  
|||||

Db 420 RNLEKNTKQBIILAALAKGCSFLPDPYQKQCDQFVAEYEPVLIELVEVMDPSFVCLKIG 479  
|||||

Qy 475 ACPSAHKPLLGTEKCIWGPSYWCNTETATAACNAVEHCKRHVWN 518  
|||||

Db 480 ACPSAHKPLLGTEKCIWGPSYWCNTETATAACNAVEHCKRHVWN 523  
|||||

RESULT 5

US-08-483-146A-2

; Sequence 2, Application US/08483146A

; Patent No. 5696080

; GENERAL INFORMATION:

; APPLICANT: O'Brien, John S.

; APPLICANT: Kishimoto, Yasuo

; TITLE OF INVENTION: PHARMACEUTICAL COMPOSITIONS

; TITLE OF INVENTION: COMPRISING PROSAPOSIN AND NEUROTROPHIC PEPTIDES DERIVED THEREFROM

; NUMBER OF SEQUENCES: 11

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Knobbe, Martens, Olson and Bear

; STREET: 620 Newport Center Blvd. 16th Floor

; CITY: Newport Beach

; STATE: CA

; COUNTRY: USA

; ZIP: 92660

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Diskette

; COMPUTER: IBM Compatible

; OPERATING SYSTEM: DOS

; SOFTWARE: FastSeq for Windows Version 2.0

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/483,146A

; FILING DATE: 07-JUN-1995

; CLASSIFICATION: 514

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER:

; FILING DATE:

; ATTORNEY/AGENT INFORMATION:

; NAME: Israelsen, Ned A.

; REGISTRATION NUMBER: 29,655

; REFERENCE/DOCKET NUMBER: MYELOS.002DV1

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: 619-235-8550

; TELEFAX: 619-235-0176

; TELEX:

; INFORMATION FOR SEQ ID NO: 2:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 523 amino acids

; TYPE: amino acid

; STRANDEDNESS: single

; TOPOLOGY: linear

; MOLECULE TYPE: peptide

; FRAGMENT TYPE: N-terminal

US-08-483-146A-2

Query Match 36.1%; Score 187; DB 1; Length 523;

Best Local Similarity 100.0%; Pred. No. 5.2e-172;

Matches 187; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 LASLLGAALAGPVLGLKECTRGSAVWCQNVKATSDCGAVKHCLQTVWNKPTVKSLPCDIC 60  
|||||

Db 7 LASLLGAALAGPVLGLKECTRGSAVWCQNVKATSDCGAVKHCLQTVWNKPTVKSLPCDIC 66  
|||||

Qy 61 KDVVTAAGDMLKONATEEILVYLEKTCDWLPKPNMSASCKEIVDSYLPVILDIKGEWS 120  
|||||

Db 67 KDVVTAAGDMLKONATEEILVYLEKTCDWLPKPNMSASCKEIVDSYLPVILDIKGEWS 126  
|||||

Qy 121 RPBGVCSALMLCESLQKHLAELNHQKLESNKIPELDMTEVVAFFMANIPILLIYPQDGP 180  
|||||

Db 127 RPBGVCSALMLCESLQKHLAELNHQKLESNKIPELDMTEVVAFFMANIPILLIYPQDGP 186  
|||||

Qy 181 SKPOPKD 187  
|||||

Db 187 SKPOPKD 193  
|||||

## RESULT 6

US-08-484-594A-2  
; Sequence 2, Application US/08484594A  
; Patent No. 5714459  
; GENERAL INFORMATION:  
; APPLICANT: O'Brien, John S.  
; APPLICANT: Kishimoto, Yasuo  
; TITLE OF INVENTION: PHARMACEUTICAL COMPOSITIONS  
; TITLE OF INVENTION: COMPRISING PROSAPOSIN AND NEUROTROPHIC PEPTIDES DERIVED  
; FROM THEREFROM  
; NUMBER OF SEQUENCES: 11  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Knobbe, Martens, Olson and Bear  
; STREET: 620 Newport Center Drive, Sixteenth Floor  
; CITY: Newport Beach  
; STATE: CA  
; COUNTRY: USA  
; ZIP: 92660  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette  
; COMPUTER: IBM Compatible  
; OPERATING SYSTEM: DOS  
; SOFTWARE: FastSeq for Windows Version 2.0  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/484,594A  
; FILING DATE: 07-JUN-1995  
; CLASSIFICATION: 514  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/100,247  
; FILING DATE: 30-JUL-1993  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Israelsen, Ned A  
; REGISTRATION NUMBER: 29,655  
; REFERENCE/DOCKET NUMBER: MYELOS, 002DV2  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 619-235-8550  
; TELEFAX: 619-235-0176  
; TELEX:  
; INFORMATION FOR SEQ ID NO: 2:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 523 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: Peptide  
; FRAGMENT TYPE: N-terminal  
US-08-484-594A-2

Query Match 36.1%; Score 187; DB 1; Length 523;  
Best Local Similarity 100.0%; Pred. No. 5.2e-172;  
Matches 187; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	1	LASLLGALAGPVLGLKECTRGSAVWCQNVKTASDCGAVKHCLQTVWNKPTVKSLPCDIC	60
Db	7	LASLLGALAGPVLGLKECTRGSAVWCQNVKTASDCGAVKHCLQTVWNKPTVKSLPCDIC	66
QY	61	KDVVTAAGDMLKONATEEEILVYLEKTCDWLPKPNMSASCKEIVDSYLPVILDIKEMS	120
Db	67	KDVVTAAGDMLKONATEEEILVYLEKTCDWLPKPNMSASCKEIVDSYLPVILDIKEMS	126
QY	121	RPGEVCSALNLCESLQKHLAELNHQKLESNKIPELDMTEVAVPFWANIPLLLYPDGGR	180
Db	127	RPGEVCSALNLCESLQKHLAELNHQKLESNKIPELDMTEVAVPFWANIPLLLYPDGGR	186
QY	181	SKPQPKD	187
Db	187	SKPQPKD	193

## RESULT 7

US-09-076-258A-2  
; Sequence 2, Application US/09076258A  
; Patent No. 6559124  
; GENERAL INFORMATION:  
; APPLICANT: O'Brien, John S.  
; APPLICANT: Kishimoto, Yasuo  
; TITLE OF INVENTION: PHARMACEUTICAL COMPOSITIONS  
; TITLE OF INVENTION: COMPRISING PROSAPOSIN AND NEUROTROPHIC PEPTIDES DERIVED  
; FROM THEREFROM  
; NUMBER OF SEQUENCES: 11  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Knobbe, Martens, Olson and Bear  
; STREET: 620 Newport Center Drive, Sixteenth Floor  
; CITY: Newport Beach  
; STATE: CA  
; COUNTRY: USA  
; ZIP: 92660  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette  
; COMPUTER: IBM Compatible  
; OPERATING SYSTEM: DOS  
; SOFTWARE: FastSeq for Windows Version 2.0  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/076,258A  
; FILING DATE: 07-JUN-1995  
; CLASSIFICATION: 514  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/483,146  
; FILING DATE: 07-JUN-1995  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Israelsen, Ned A  
; REGISTRATION NUMBER: 29,655  
; REFERENCE/DOCKET NUMBER: MYELOS, 2DV1C2  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 619-235-8550  
; TELEFAX: 619-235-0176  
; TELEX:  
; INFORMATION FOR SEQ ID NO: 2:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 523 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: Peptide  
; FRAGMENT TYPE: N-terminal  
US-09-076-258A-2

Query Match 36.1%; Score 187; DB 4; Length 523;  
Best Local Similarity 100.0%; Pred. No. 5.2e-172;  
Matches 187; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	1	LASLLGALAGPVLGLKECTRGSAVWCQNVKTASDCGAVKHCLQTVWNKPTVKSLPCDIC	60
Db	7	LASLLGALAGPVLGLKECTRGSAVWCQNVKTASDCGAVKHCLQTVWNKPTVKSLPCDIC	66
QY	61	KDVVTAAGDMLKONATEEEILVYLEKTCDWLPKPNMSASCKEIVDSYLPVILDIKEMS	120
Db	67	KDVVTAAGDMLKONATEEEILVYLEKTCDWLPKPNMSASCKEIVDSYLPVILDIKEMS	126
QY	121	RPGEVCSALNLCESLQKHLAELNHQKLESNKIPELDMTEVAVPFWANIPLLLYPDGGR	180
Db	127	RPGEVCSALNLCESLQKHLAELNHQKLESNKIPELDMTEVAVPFWANIPLLLYPDGGR	186
QY	181	SKPQPKD	187
Db	187	SKPQPKD	193

## RESULT 8

US-09-352-548-2  
; Sequence 2, Application US/09352548  
; Patent No. 6500431

GENERAL INFORMATION:  
APPLICANT: Gill, Parkash S.  
TITLE OF INVENTION: No. 6500431el Inhibitors of Angiogenesis and Tumor Growth  
FILE REFERENCE: 017986-000410US  
CURRENT APPLICATION NUMBER: US/09/352,548  
CURRENT FILING DATE: 1999-07-12  
EARLIER APPLICATION NUMBER: US 60/092,647  
EARLIER FILING DATE: 1998-07-13  
NUMBER OF SEQ ID NOS: 59  
SOFTWARE: PatentIn Ver. 2.1  
SEQ ID NO 2  
LENGTH: 81  
TYPE: PRT  
ORGANISM: Homo sapiens  
FEATURE:  
OTHER INFORMATION: Saposin B  
US-09-352-548-2

Query Match 15.6%; Score 81; DB 4; Length 81;  
Best Local Similarity 100.0%; Pred. No. 1.8e-70;  
Matches 81; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 189 GDVQCDCIQWTDIQTAVRTNSTFVQALVHVHKECDRLGPGMADICKNYISQYSEIAIQ 248  
Db 1 GDVQCDCIQWTDIQTAVRTNSTFVQALVHVHKECDRLGPGMADICKNYISQYSEIAIQ 60  
Qy 249 MMMHMQKEICALVGFCDVEK 269  
Db 61 MMMHMQKEICALVGFCDVEK 81

RESULT 9  
US-08-100-247-3  
Sequence 3, Application US/08100247  
Patent No. 5571787

GENERAL INFORMATION:  
APPLICANT: O'BRIEN, JOHN S.  
TITLE OF INVENTION: PROSAPOSIN AS A NEUROTROPHIC FACTOR  
NUMBER OF SEQUENCES: 5  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: KNOBBE, MARTENS, OLSON AND BEAR  
STREET: 620 NEWPORT CENTER DRIVE SIXTEENTH FLOOR  
CITY: NEWPORT BEACH  
STATE: CA  
COUNTRY: USA  
ZIP: 92660

COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/100,247  
FILING DATE: 19930730  
CLASSIFICATION: 514

ATTORNEY/AGENT INFORMATION:  
NAME: Israelsen, Ned A.  
REGISTRATION NUMBER: 29,655  
REFERENCE/DOCKET NUMBER: O'BRIEN.002A  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 619-235-8550  
TELEFAX: 619-235-0176

INFORMATION FOR SEQ ID NO: 3:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 80 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
HYPOTHETICAL: NO  
ANTI-SENSE: NO

FRAGMENT TYPE: internal  
IMMEDIATE SOURCE:  
CLONE: SAPOSIN C  
US-08-100-247-3

Query Match 15.4%; Score 80; DB 1; Length 80;  
Best Local Similarity 100.0%; Pred. No. 1.7e-69;  
Matches 80; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 305 SDVYCEVCEFLVKEVTKLIDNNKTEKILDAFDKMCCKLPKSLSECEQVVDVTVGSSILS 364  
Db 1 SDVYCEVCEFLVKEVTKLIDNNKTEKILDAFDKMCCKLPKSLSECEQVVDVTVGSSILS 60  
Qy 365 ILLEEVSPELVCSMLHLCSG 384  
Db 61 ILLEEVSPELVCSMLHLCSG 80

RESULT 10

US-08-483-146A-3  
Sequence 3, Application US/08483146A  
Patent No. 5696080

GENERAL INFORMATION:  
APPLICANT: O'Brien, John S.  
APPLICANT: Kishimoto, Yasuo  
TITLE OF INVENTION: PHARMACEUTICAL COMPOSITIONS  
TITLE OF INVENTION: COMPRISING PROSAPOSIN AND NEUROTROPHIC PEPTIDES DERIVED  
TITLE OF INVENTION: THEREFROM  
NUMBER OF SEQUENCES: 11  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Knobbe, Martens, Olson and Bear  
STREET: 620 Newport Center Blvd. 16th Floor  
CITY: Newport Beach  
STATE: CA  
COUNTRY: USA  
ZIP: 92660

COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: DOS  
SOFTWARE: FastSeq for Windows Version 2.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/483,146A  
FILING DATE: 07-JUN-1995  
CLASSIFICATION: 514

PRIOR APPLICATION DATA:  
APPLICATION NUMBER:  
FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: Israelsen, Ned A.  
REGISTRATION NUMBER: 29,655  
REFERENCE/DOCKET NUMBER: MYELOS.002DV1  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 619-235-8550  
TELEFAX: 619-235-0176  
TELEX:  
INFORMATION FOR SEQ ID NO: 3:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 80 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
FRAGMENT TYPE: internal

US-08-483-146A-3  
Query Match 15.4%; Score 80; DB 1; Length 80;  
Best Local Similarity 100.0%; Pred. No. 1.7e-69;  
Matches 80; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 305 SDVYCEVCEFLVKEVTKLIDNNKTEKILDAFDKMCCKLPKSLSECEQVVDVTVGSSILS 364  
Db 1 SDVYCEVCEFLVKEVTKLIDNNKTEKILDAFDKMCCKLPKSLSECEQVVDVTVGSSILS 60

QY 365 ILLEEVSPELVCSMLHLCSG 384  
Db 61 ILLEEVSPELVCSMLHLCSG 80

## RESULT 11

US-08-232-513A-4  
; Sequence 4, Application US/08232513A  
; Patent No. 5700909  
; GENERAL INFORMATION:  
; APPLICANT: O'Brien, John S.  
; TITLE OF INVENTION: Prosaposin and Cytokine-Derived Peptides  
; NUMBER OF SEQUENCES: 20  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Campbell & Flores LLP  
; STREET: 4370 La Jolla Village Drive, Suite 700  
; CITY: San Diego  
; STATE: California  
; COUNTRY: USA  
; ZIP: 92122  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/232,513A  
; FILING DATE: 21-APR-1994  
; CLASSIFICATION: 514  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/100,247  
; FILING DATE: 30-JUL-1993  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Campbell, Cathryn A.  
; REGISTRATION NUMBER: 31,815  
; REFERENCE/DOCKET NUMBER: P-UD 1643  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (619) 535-8949  
; TELEFAX: (619) 535-9001  
; INFORMATION FOR SEQ ID NO: 4:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 80 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
; FEATURE:  
; NAME/KEY: Peptide  
; LOCATION: 1..80  
; OTHER INFORMATION: /label= Saposin\_C  
US-08-232-513A-4

Query Match 15.4%; Score 80; DB 1; Length 80;  
Best Local Similarity 100.0%; Pred. No. 1.7e-69;  
Matches 80; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 305 SDVYCEVCEFLVKEVTKLIDNNKTEKILDAFDMCKSLPKSLSEECQEVVDITYGSSILS 364  
Db 1 SDVYCEVCEFLVKEVTKLIDNNKTEKILDAFDMCKSLPKSLSEECQEVVDITYGSSILS 60  
QY 365 ILLEEVSPELVCSMLHLCSG 384  
Db 61 ILLEEVSPELVCSMLHLCSG 80

## RESULT 12

US-08-484-594A-3  
; Sequence 3, Application US/08484594A  
; Patent No. 5714459  
; GENERAL INFORMATION:  
; APPLICANT: O'Brien, John S.  
; APPLICANT: Kishimoto, Yasuo

; TITLE OF INVENTION: USE OF PROSAPOSIN AND NEUTROTROPIC PEPTIDES  
; TITLE OF INVENTION: DERIVED THEREFROM  
; NUMBER OF SEQUENCES: 11  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Knobbe, Martens, Olson and Bear  
; STREET: 620 Newport Center Drive, Sixteenth Floor  
; CITY: Newport Beach  
; STATE: CA  
; COUNTRY: USA  
; ZIP: 92660  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette  
; COMPUTER: IBM Compatible  
; OPERATING SYSTEM: DOS  
; SOFTWARE: FastSeq for Windows Version 2.0  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/484,594A  
; FILING DATE: 07-JUN-1995  
; CLASSIFICATION: 514  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/100,247  
; FILING DATE: 30-JUL-1993  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Israelsen, Ned A.  
; REGISTRATION NUMBER: 29,655  
; REFERENCE/DOCKET NUMBER: MYELOS.002DV2  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 619-235-8850  
; TELEFAX: 619-235-0176  
; TELEX:  
; INFORMATION FOR SEQ ID NO: 3:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 80 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; FRAGMENT TYPE: internal  
US-08-484-594A-3  
Query Match 15.4%; Score 80; DB 1; Length 80;  
Best Local Similarity 100.0%; Pred. No. 1.7e-69;  
Matches 80; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 305 SDVYCEVCEFLVKEVTKLIDNNKTEKILDAFDMCKSLPKSLSEECQEVVDITYGSSILS 364  
Db 1 SDVYCEVCEFLVKEVTKLIDNNKTEKILDAFDMCKSLPKSLSEECQEVVDITYGSSILS 60  
QY 365 ILLEEVSPELVCSMLHLCSG 384  
Db 61 ILLEEVSPELVCSMLHLCSG 80

## RESULT 13

US-09-076-258A-3  
; Sequence 3, Application US/09076258A  
; Patent No. 6559124  
; GENERAL INFORMATION:  
; APPLICANT: O'Brien, John S.  
; APPLICANT: Kishimoto, Yasuo  
; TITLE OF INVENTION: PHARMACEUTICAL COMPOSITIONS  
; TITLE OF INVENTION: COMPRISING PROSAPOSIN AND NEUTROTROPIC PEPTIDES DERIVED  
; NUMBER OF SEQUENCES: 11  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Knobbe, Martens, Olson and Bear  
; STREET: 620 Newport Center Blvd. 16th Floor  
; CITY: Newport Beach  
; STATE: CA  
; COUNTRY: USA  
; ZIP: 92660  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette



COMPUTER: IBM Compatible  
OPERATING SYSTEM: DOS  
SOFTWARE: FastSeq for Windows Version 2.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/076,258A  
FILING DATE:

## CLASSIFICATION:

PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/958,970  
FILING DATE: 28-OCT-97  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/483,146  
FILING DATE: 07-JUN-1995  
APPLICATION NUMBER: 08/100,247  
FILING DATE: 30-JUL-1993

## ATTORNEY/AGENT INFORMATION:

NAME: Israelsen, Ned A.  
REGISTRATION NUMBER: 29,655  
REFERENCE/DOCKET NUMBER: MYELOS.2DV1C2  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 619-235-8550  
TELEFAX: 619-235-0176  
TELEX:

## INFORMATION FOR SEQ ID NO: 3:

SEQUENCE CHARACTERISTICS:  
LENGTH: 80 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
FRAGMENT TYPE: internal

US-09-076-258A-3

Query Match 15.4%; Score 80; DB 4; Length 80;  
Best Local Similarity 100.0%; Pred. No. 1.7e-69;  
Matches 80; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 305 SDVYCEVCEFLVKEVTKLIDNNKTEKEILDADFDMCKSLPKSLSEECQEVVDVTYGSILS 364  
Db 1 SDVYCEVCEFLVKEVTKLIDNNKTEKEILDADFDMCKSLPKSLSEECQEVVDVTYGSILS 60  
Qy 365 ILLEEVSPELVCSMLHLCSG 384  
Db 61 ILLEEVSPELVCSMLHLCSG 80

## RESULT 14

US-08-756-031-3  
Sequence 3, Application US/08756031  
Patent No. 6590074

## GENERAL INFORMATION:

APPLICANT: O'BRIEN, JOHN S.  
APPLICANT: KISHIMOTO, YASUO  
TITLE OF INVENTION: PROSAPOSIN AS A NEUROTROPHIC FACTOR  
NUMBER OF SEQUENCES: 5  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: KNOBBE, MARTENS, OLSON AND BEAR  
STREET: 620 NEWPORT CENTER DRIVE SIXTEENTH FLOOR  
CITY: NEWPORT BEACH  
STATE: CA  
COUNTRY: USA  
ZIP: 92660

## COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent In Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/756,031  
FILING DATE: 26-NOV-1996  
CLASSIFICATION: 514  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/100,247

FILING DATE: 30-JUL-1993  
ATTORNEY/AGENT INFORMATION:  
NAME: Israelsen, Ned A.  
REGISTRATION NUMBER: 29,655  
REFERENCE/DOCKET NUMBER: OBRIEN.002A  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 619-235-8550  
TELEFAX: 619-235-0176

## INFORMATION FOR SEQ ID NO: 3:

SEQUENCE CHARACTERISTICS:  
LENGTH: 80 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
HYPOTHETICAL: NO  
ANTI-SENSE: NO  
FRAGMENT TYPE: internal  
IMMEDIATE SOURCE:  
CLONE: SAPOSIN C

US-08-756-031-3

Query Match 15.4%; Score 80; DB 4; Length 80;

Best Local Similarity 100.0%; Pred. No. 1.7e-69;

Matches 80; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

---

Qy 305 SDVYCEVCEFLVKEVTKLIDNNKTEKEILDADFDMCKSLPKSLSEECQEVVDVTYGSILS 364  
Db 1 SDVYCEVCEFLVKEVTKLIDNNKTEKEILDADFDMCKSLPKSLSEECQEVVDVTYGSILS 60

Qy 365 ILLEEVSPELVCSMLHLCSG 384

Db 61 ILLEEVSPELVCSMLHLCSG 80

RESULT 15

US-08-584-671-15

Sequence 15, Application US/08584671

Patent No. 5910568

GENERAL INFORMATION:

APPLICANT: HAMMERSTEDT, ROY H, BARBATO, GUY F,

APPLICANT: CRAMER, PALMER

TITLE OF INVENTION: MOLECULE INVOLVED IN BINDING OF SPERM

TITLE OF INVENTION: TO EGG SURFACES AND PROCEDURES FOR USE OF THIS MOLECULE

TITLE OF INVENTION: TO ENHANCE OR DECREASE POTENTIAL FERTILITY

NUMBER OF SEQUENCES: 16

CORRESPONDENCE ADDRESS:

ADDRESSEE: INTELLECTUAL PROPERTY OFFICE, THE PENNSYLVANIA

ADDRESSEE: STATE UNIVERSITY

STREET: 113 TECHNOLOGY CENTER

CITY: UNIVERSITY PARK

STATE: PENNSYLVANIA

COUNTRY: UNITED STATES OF AMERICA

ZIP: 16802-7000

COMPUTER READABLE FORM:

MEDIUM TYPE: FLOPPY DISK

COMPUTER: NEC-286

OPERATING SYSTEM: DOS

SOFTWARE: WORDPERFECT 5.1

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/584,671

FILING DATE:

CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:

NAME: MONAHAN, THOMAS J

REGISTRATION NUMBER: 29835

REFERENCE/DOCKET NUMBER:

TELECOMMUNICATION INFORMATION:

TELEPHONE: 814-865-6277

TELEFAX: 814-865-3591

INFORMATION FOR SEQ ID NO: 15:

SEQUENCE CHARACTERISTICS:

LENGTH: 80

;  
; TYPE: AMINO ACID  
; STRANDEDNESS: SINGLE  
; TOPOLOGY: UNKNOWN  
US-08-584-671-15

Query Match 14.5%; Score 75; DB 2; Length 80;  
Best Local Similarity 100.0%; Pred. No. 1.1e-64;  
Matches 75; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 128 ALNLCESLQKHLAELNHQKLESNKIPELDMTEVVVAPFMANIPLLLYPDGGRSKPKD 187  
Db |||||  
6 ALNLCESLQKHLAELNHQKLESNKIPELDMTEVVVAPFMANIPLLLYPDGGRSKPKD 65  
QY 188 NGDVCODCIQWTDI 202  
Db |||||  
66 NGDVCODCIQWTDI 80

Search completed: May 5, 2004, 13:35:19  
Job time : 17.405 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 5, 2004, 13:33:49 ; Search time 41.7582 Seconds  
(without alignments)  
3438.471 Million cell updates/sec

Title: US-09-743-684A-1\_COPY\_7\_524

Perfect score: 518  
Sequence: 1 LASLLGALAGPVILGKECT.....NTETAACQNAVEHCKRHVWN 518

Scoring table: OLIGO  
Gapop 60.0 , Gapext 60.0

Searched: 1138120 seqs, 277189581 residues

Word size : 0

Total number of hits satisfying chosen parameters: 1138120

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

- Database : Published Applications AA:\*
- 1: /cgn2\_6/ptodata/2/pubpaa/PCT\_NEW\_PUB.pdb.\*
  - 2: /cgn2\_6/ptodata/2/pubpaa/PCT\_NEW\_PUB.pdb.\*
  - 3: /cgn2\_6/ptodata/2/pubpaa/US06\_NEW\_PUB.pdb.\*
  - 4: /cgn2\_6/ptodata/2/pubpaa/US06\_PUBCOMB.pdb.\*
  - 5: /cgn2\_6/ptodata/2/pubpaa/US07\_NEW\_PUB.pdb.\*
  - 6: /cgn2\_6/ptodata/2/pubpaa/PCTUS\_PUBCOMB.pdb.\*
  - 7: /cgn2\_6/ptodata/2/pubpaa/US08\_NEW\_PUB.pdb.\*
  - 8: /cgn2\_6/ptodata/2/pubpaa/US08\_PUBCOMB.pdb.\*
  - 9: /cgn2\_6/ptodata/2/pubpaa/US09A\_PUBCOMB.pdb.\*
  - 10: /cgn2\_6/ptodata/2/pubpaa/US09B\_PUBCOMB.pdb.\*
  - 11: /cgn2\_6/ptodata/2/pubpaa/US09C\_PUBCOMB.pdb.\*
  - 12: /cgn2\_6/ptodata/2/pubpaa/US09\_NEW\_PUB.pdb.\*
  - 13: /cgn2\_6/ptodata/2/pubpaa/US10A\_PUBCOMB.pdb.\*
  - 14: /cgn2\_6/ptodata/2/pubpaa/US10B\_PUBCOMB.pdb.\*
  - 15: /cgn2\_6/ptodata/2/pubpaa/US10C\_PUBCOMB.pdb.\*
  - 16: /cgn2\_6/ptodata/2/pubpaa/US10\_NEW\_PUB.pdb.\*
  - 17: /cgn2\_6/ptodata/2/pubpaa/US60\_NEW\_PUB.pdb.\*
  - 18: /cgn2\_6/ptodata/2/pubpaa/US60\_PUBCOMB.pdb.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	518	100.0	524	9	US-09-870-759-60
2	518	100.0	524	10	US-09-751-708A-60
3	518	100.0	524	12	US-10-267-502-386
4	278	53.7	527	9	US-09-870-759-61
5	278	53.7	527	10	US-09-751-708A-61
6	278	53.7	527	14	US-10-060-036-73
7	254	49.0	479	10	US-09-978-418-40
8	209	40.3	209	14	US-10-043-487-340
9	187	36.1	523	9	US-09-767-007A-2
10	80	15.4	80	9	US-09-767-007A-3
11	80	15.4	80	9	US-09-753-126-3
12	80	15.4	80	15	US-10-330-697-3
13	80	15.4	592	9	US-09-753-126-4
14	80	15.4	592	15	US-10-330-697-4
15	40	7.7	40	10	US-09-780-438A-1

16	38	7.3	38	10	US-09-780-438A-2	Sequence 2, Appli
17	156	12	156	12	US-09-925-298-644	Sequence 644, App
18	24	4.6	24	14	US-10-102-806-644	Sequence 644, App
19	22	4.2	22	9	US-09-767-007A-1	Sequence 1, Appli
20	22	4.2	22	9	US-09-767-007A-11	Sequence 11, Appli
21	22	4.2	22	9	US-09-957-143-1	Sequence 1, Appli
22	20	3.9	554	14	US-10-205-194-176	Sequence 176, App
23	17	3.3	18	9	US-09-767-007A-5	Sequence 5, Appli
24	17	3.3	18	9	US-09-957-143-2	Sequence 2, Appli
25	15	2.9	15	9	US-09-802-617-2	Sequence 2, Appli
26	15	2.9	15	14	US-10-293-819-2	Sequence 2, Appli
27	12	2.3	12	9	US-09-802-617-1	Sequence 1, Appli
28	12	2.3	12	9	US-09-957-143-3	Sequence 3, Appli
29	12	2.3	12	14	US-10-293-819-1	Sequence 1, Appli
30	11	2.1	11	10	US-09-791-393-182	Sequence 182, App
31	11	2.1	11	10	US-09-791-389-182	Sequence 182, App
32	11	2.1	11	14	US-10-281-478-126	Sequence 126, App
33	10	1.9	15	14	US-10-281-478-74	Sequence 74, Appli
34	10	1.9	20	14	US-10-281-478-77	Sequence 77, Appli
35	9	1.7	14	14	US-10-281-478-75	Sequence 129, App
36	9	1.7	17	14	US-10-281-478-129	Sequence 4623, Ap
37	8	1.5	286	15	US-10-108-260A-4623	Sequence 279814, A
38	8	1.5	354	12	US-10-424-599-279814	Sequence 46329, A
39	8	1.5	361	12	US-10-425-114-46329	Sequence 72368, A
40	8	1.5	363	12	US-10-425-114-72368	Sequence 10328, A
41	8	1.5	655	9	US-09-815-342-10328	Sequence 56714, A
42	8	1.5	655	12	US-10-282-122A-56714	Sequence 34, Appli
43	8	1.5	1896	12	US-10-312-352-34	Sequence 35398, A
44	7	1.4	46	9	US-09-864-761-35398	Sequence 211534, A
45	7	1.4	50	12	US-10-424-599-211534	

ALIGNMENTS

RESULT 1  
US-09-870-759-60  
; Sequence 60, Application US/09870759  
; Patent No. US20020177551A1  
; GENERAL INFORMATION:  
; APPLICANT: TERMAN, David S  
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATMENT OF NEOPLASTIC DISEASE  
; FILE REFERENCE: 870759  
; CURRENT APPLICATION NUMBER: US/09/870,759  
; CURRENT FILING DATE: 2002-01-14  
; PRIOR APPLICATION NUMBER: US 60/208,128  
; PRIOR FILING DATE: 2000-05-30  
; NUMBER OF SEQ ID NOS: 166  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 60  
; LENGTH: 524  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-870-759-60

Query Match	100.0%	Score 518;	DB 9;	Length 524;
Best Local Similarity	100.0%;	Pred. No. 0;		
Matches 518;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
Qy	1	LASLLGALAGPVILGKECTRGSAVWCNQVKTASDCGAVKHCLQTVWNKPTVKSLPCDIC	60	
Db	7	LASLLGALAGPVILGKECTRGSAVWCNQVKTASDCGAVKHCLQTVWNKPTVKSLPCDIC	66	
Qy	61	KDVVTAAGMDLKNATTEETILVLEKTCMDLPKNMSASCKEIVDSYLPVILDIKIGMS	120	
Db	67	KDVVTAAGMDLKNATTEETILVLEKTCMDLPKNMSASCKEIVDSYLPVILDIKIGMS	126	
Qy	121	RPGEVCSALNLCESLQKHAEHNLHQLKESNIPKPELDMTEVAVPFMANIPILLYPQDQPR	180	
Db	127	RPGEVCSALNLCESLQKHAEHNLHQLKESNIPKPELDMTEVAVPFMANIPILLYPQDQPR	186	
Qy	181	SKPQKONGDVCPQCIQMTDITQAVRTNSTFVQALVEHVKCECDRLGPGMADICKNYS	240	

Db 187 SKPOKONGDVQCDCIQWVTDIQTAVRTNSTFVQALVEHVKECDRLGPGMADICKNYIS 246  
QY 241 QYSEIAIQMMHMQPKEICALVGFCDEVKEMPQTLVPAKASKNVI PALELVEPIKKHE 300  
Db 247 QYSEIAIQMMHMQPKEICALVGFCDEVKEMPQTLVPAKASKNVI PALELVEPIKKHE 306  
QY 301 VPAKSDVYCEVCEFLVKEVTKLIDNNKTEKILDAFDKMSKLPKSLSECEQVVDITYGS 360  
Db 307 VPAKSDVYCEVCEFLVKEVTKLIDNNKTEKILDAFDKMSKLPKSLSECEQVVDITYGS 366  
QY 361 SILSILLEEVSPELVCSMLHLCSGTRLPALTTHVTQPKDGGFCVCKKLVGLDRNLKN 420  
Db 367 SILSILLEEVSPELVCSMLHLCSGTRLPALTTHVTQPKDGGFCVCKKLVGLDRNLKN 426  
QY 421 STKOEILAALEKGCSEFLPDYQKQCQDFVAEYEPVLIEILVEVMDPSFVCLKIGACPSAH 480  
Db 427 STKOEILAALEKGCSEFLPDYQKQCQDFVAEYEPVLIEILVEVMDPSFVCLKIGACPSAH 486  
QY 481 KPLLGTKEKCIWGPSYWCQNTETAACQNAVEHCKRHWVN 518  
Db 487 KPLLGTKEKCIWGPSYWCQNTETAACQNAVEHCKRHWVN 524

RESULT 2  
US-09-751-708A-60  
; Sequence 60, Application US/09751708A  
; Publication No. US20030157113A1  
; GENERAL INFORMATION:  
; APPLICANT: TERMAN, David S  
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATMENT OF NEOPLASTIC DISEASE  
; FILE REFERENCE: 751708  
; CURRENT APPLICATION NUMBER: US/09/751,708A  
; CURRENT FILING DATE: 2002-10-15  
; PRIOR APPLICATION NUMBER: US 60/173,371  
; PRIOR FILING DATE: 1999-12-28  
; NUMBER OF SEQ ID NOS: 166  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 60  
; LENGTH: 524  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-751-708A-60

Query Match 100.0%; Score 518; DB 10; Length 524;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 518; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LASLLGAALAGPVLGLKECTRGSAVWCQNVKTASDCGAVKHCLQTVWNKPTVKSLLPCDIC 60  
Db 7 LASLLGAALAGPVLGLKECTRGSAVWCQNVKTASDCGAVKHCLQTVWNKPTVKSLLPCDIC 66  
QY 61 KDVTAAAGDMLKDNATEEELVLEKTCDWLPKPNMSASCKEIVDSYLPVILDIKGENS 120  
Db 67 KDVTAAAGDMLKDNATEEELVLEKTCDWLPKPNMSASCKEIVDSYLPVILDIKGENS 126  
QY 121 RPGEVCSALNLCESLQKHLAELNHQKLESNKIPELDMTEVAVPAMANIPLLLYPQDGP 180  
Db 127 RPGEVCSALNLCESLQKHLAELNHQKLESNKIPELDMTEVAVPAMANIPLLLYPQDGP 186  
QY 181 SKPOKONGDVQCDCIQWVTDIQTAVRTNSTFVQALVEHVKECDRLGPGMADICKNYIS 240  
Db 187 SKPOKONGDVQCDCIQWVTDIQTAVRTNSTFVQALVEHVKECDRLGPGMADICKNYIS 246  
QY 241 QYSEIAIQMMHMQPKEICALVGFCDEVKEMPQTLVPAKASKNVI PALELVEPIKKHE 300  
Db 247 QYSEIAIQMMHMQPKEICALVGFCDEVKEMPQTLVPAKASKNVI PALELVEPIKKHE 306  
QY 301 VPAKSDVYCEVCEFLVKEVTKLIDNNKTEKILDAFDKMSKLPKSLSECEQVVDITYGS 360  
Db 307 VPAKSDVYCEVCEFLVKEVTKLIDNNKTEKILDAFDKMSKLPKSLSECEQVVDITYGS 366  
QY 361 SILSILLEEVSPELVCSMLHLCSGTRLPALTTHVTQPKDGGFCVCKKLVGLDRNLKN 420  
Db 367 SILSILLEEVSPELVCSMLHLCSGTRLPALTTHVTQPKDGGFCVCKKLVGLDRNLKN 426  
QY 421 STKOEILAALEKGCSEFLPDYQKQCQDFVAEYEPVLIEILVEVMDPSFVCLKIGACPSAH 480  
Db 427 STKOEILAALEKGCSEFLPDYQKQCQDFVAEYEPVLIEILVEVMDPSFVCLKIGACPSAH 486  
QY 481 KPLLGTKEKCIWGPSYWCQNTETAACQNAVEHCKRHWVN 518  
Db 487 KPLLGTKEKCIWGPSYWCQNTETAACQNAVEHCKRHWVN 524

Db 367 SILSILLEEVSPELVCSMLHLCSGTRLPALTTHVTQPKDGGFCVCKKLVGLDRNLKN 426  
QY 421 STKOEILAALEKGCSEFLPDYQKQCQDFVAEYEPVLIEILVEVMDPSFVCLKIGACPSAH 480  
Db 427 STKOEILAALEKGCSEFLPDYQKQCQDFVAEYEPVLIEILVEVMDPSFVCLKIGACPSAH 486  
QY 481 KPLLGTKEKCIWGPSYWCQNTETAACQNAVEHCKRHWVN 518  
Db 487 KPLLGTKEKCIWGPSYWCQNTETAACQNAVEHCKRHWVN 524

RESULT 3  
US-10-267-502-386  
; Sequence 386, Application US/10267502  
; Publication No. US20040071700A1  
; GENERAL INFORMATION:  
; APPLICANT: Kim, Jaeseob  
; APPLICANT: Galant, Ron  
; TITLE OF INVENTION: Obesity Linked Genes  
; FILE REFERENCE: LSD-07416  
; CURRENT APPLICATION NUMBER: US/10/267,502  
; CURRENT FILING DATE: 2003-01-27  
; NUMBER OF SEQ ID NOS: 439  
; SOFTWARE: PatentIn version 3.2  
; SEQ ID NO 386  
; LENGTH: 524  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-267-502-386

Query Match 100.0%; Score 518; DB 12; Length 524;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 518; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LASLLGAALAGPVLGLKECTRGSAVWCQNVKTASDCGAVKHCLQTVWNKPTVKSLLPCDIC 60  
Db 7 LASLLGAALAGPVLGLKECTRGSAVWCQNVKTASDCGAVKHCLQTVWNKPTVKSLLPCDIC 66  
QY 61 KDVTAAAGDMLKDNATEEELVLEKTCDWLPKPNMSASCKEIVDSYLPVILDIKGENS 120  
Db 67 KDVTAAAGDMLKDNATEEELVLEKTCDWLPKPNMSASCKEIVDSYLPVILDIKGENS 126  
QY 121 RPGEVCSALNLCESLQKHLAELNHQKLESNKIPELDMTEVAVPAMANIPLLLYPQDGP 180  
Db 127 RPGEVCSALNLCESLQKHLAELNHQKLESNKIPELDMTEVAVPAMANIPLLLYPQDGP 186  
QY 181 SKPOKONGDVQCDCIQWVTDIQTAVRTNSTFVQALVEHVKECDRLGPGMADICKNYIS 240  
Db 187 SKPOKONGDVQCDCIQWVTDIQTAVRTNSTFVQALVEHVKECDRLGPGMADICKNYIS 246  
QY 241 QYSEIAIQMMHMQPKEICALVGFCDEVKEMPQTLVPAKASKNVI PALELVEPIKKHE 300  
Db 247 QYSEIAIQMMHMQPKEICALVGFCDEVKEMPQTLVPAKASKNVI PALELVEPIKKHE 306  
QY 301 VPAKSDVYCEVCEFLVKEVTKLIDNNKTEKILDAFDKMSKLPKSLSECEQVVDITYGS 360  
Db 307 VPAKSDVYCEVCEFLVKEVTKLIDNNKTEKILDAFDKMSKLPKSLSECEQVVDITYGS 366  
QY 361 SILSILLEEVSPELVCSMLHLCSGTRLPALTTHVTQPKDGGFCVCKKLVGLDRNLKN 420  
Db 367 SILSILLEEVSPELVCSMLHLCSGTRLPALTTHVTQPKDGGFCVCKKLVGLDRNLKN 426  
QY 421 STKOEILAALEKGCSEFLPDYQKQCQDFVAEYEPVLIEILVEVMDPSFVCLKIGACPSAH 480  
Db 427 STKOEILAALEKGCSEFLPDYQKQCQDFVAEYEPVLIEILVEVMDPSFVCLKIGACPSAH 486  
QY 481 KPLLGTKEKCIWGPSYWCQNTETAACQNAVEHCKRHWVN 518  
Db 487 KPLLGTKEKCIWGPSYWCQNTETAACQNAVEHCKRHWVN 524

RESULT 4  
US-09-870-759-61

; Sequence 61, Application US/09870759  
; Patent No. US20020177551A1  
; GENERAL INFORMATION:  
; APPLICANT: TERMAN, David S  
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATMENT OF NEOPLASTIC DISEASE  
; FILE REFERENCE: 870759  
; CURRENT APPLICATION NUMBER: US/09/870,759  
; CURRENT FILING DATE: 2002-01-14  
; PRIOR APPLICATION NUMBER: US 60/208,128  
; PRIOR FILING DATE: 2000-05-30  
; NUMBER OF SEQ ID NOS: 166  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 61  
; LENGTH: 527  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-870-759-61

Query Match 53.7%; Score 278; DB 9; Length 527;  
Best Local Similarity 99.4%; Pred. No. 1.9e-263;  
Matches 518; Conservative 0; Mismatches 0; Indels 3; Gaps 1;

Qy 1 LASLLGALAGPVLGLKECTRGSAVWCONVKTASDCGAVKHCLQTVWNKPTVKSILPCDIC 60  
Db 7 LASLLGALAGPVLGLKECTRGSAVWCONVKTASDCGAVKHCLQTVWNKPTVKSILPCDIC 66

Qy 61 KDVVTAAGDMLKDNATEEEILVLEKTCDWLPKPNMSASCKEIVDSYLPVILDIKIGEMS 120  
Db 67 KDVVTAAGDMLKDNATEEEILVLEKTCDWLPKPNMSASCKEIVDSYLPVILDIKIGEMS 126

Qy 121 RPEVCSALNLCESLQKHLAELNHQKLESNKIPELDMTEVAVPFMANIPLLLYPQDGP 180  
Db 127 RPEVCSALNLCESLQKHLAELNHQKLESNKIPELDMTEVAVPFMANIPLLLYPQDGP 186

Qy 181 SKQPKDNGVCDQCIQWVTDIQTAVRTNSTFVQALVEHVKEECDRILGPGWADICKNYIS 240  
Db 187 SKQPKDNGVCDQCIQWVTDIQTAVRTNSTFVQALVEHVKEECDRILGPGWADICKNYIS 246

Qy 241 QYSEIATQMMHMM--OPKEICALVGFCDVEKEMPMOTLPAKVASKNVI PALELVEPIK 297  
Db 247 QYSEIATQMMHMMQDQPKKEICALVGFCDVEKEMPMOTLPAKVASKNVI PALELVEPIK 306

Qy 298 KHEVPAKSDVYCEVCEFLVKEVTKLIDNNKTEKILDAFDKMSKLPKSLSEECQEVVDT 357  
Db 307 KHEVPAKSDVYCEVCEFLVKEVTKLIDNNKTEKILDAFDKMSKLPKSLSEECQEVVDT 366

Qy 358 YGSSILSILLLEEVSPELVCSMLHLCSTGLPALTVHTVTPKDGGFCEVCKKLGYLDRL 417  
Db 367 YGSSILSILLLEEVSPELVCSMLHLCSTGLPALTVHTVTPKDGGFCEVCKKLGYLDRL 426

Qy 418 EKSTKOEILAALEKSGSFLPDYKQCDQFVAEYEPVLIILVEVMDPSFVCLKIGACP 477  
Db 427 EKSTKOEILAALEKSGSFLPDYKQCDQFVAEYEPVLIILVEVMDPSFVCLKIGACP 486

Qy 478 SAHKPLLGTEKICWGPSYWCQNTETAACNAVEHCKRHVWN 518  
Db 487 SAHKPLLGTEKICWGPSYWCQNTETAACNAVEHCKRHVWN 527

RESULT 5  
US-09-751-708A-61  
; Sequence 61, Application US/09751708A  
; Publication No. US20030157113A1  
; GENERAL INFORMATION:  
; APPLICANT: TERMAN, David S  
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATMENT OF NEOPLASTIC DISEASE  
; FILE REFERENCE: 751708  
; CURRENT APPLICATION NUMBER: US/09/751,708A  
; CURRENT FILING DATE: 2002-10-15  
; PRIOR APPLICATION NUMBER: US 60/173,371  
; PRIOR FILING DATE: 1999-12-28  
; NUMBER OF SEQ ID NOS: 166  
; SOFTWARE: PatentIn version 3.1

; SEQ ID NO 61  
; LENGTH: 527  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-751-708A-61

Query Match 53.7%; Score 278; DB 10; Length 527;  
Best Local Similarity 99.4%; Pred. No. 1.9e-263;  
Matches 518; Conservative 0; Mismatches 0; Indels 3; Gaps 1;

Qy 1 LASLLGALAGPVLGLKECTRGSAVWCONVKTASDCGAVKHCLQTVWNKPTVKSILPCDIC 60  
Db 7 LASLLGALAGPVLGLKECTRGSAVWCONVKTASDCGAVKHCLQTVWNKPTVKSILPCDIC 66

Qy 61 KDVVTAAGDMLKDNATEEEILVLEKTCDWLPKPNMSASCKEIVDSYLPVILDIKIGEMS 120  
Db 67 KDVVTAAGDMLKDNATEEEILVLEKTCDWLPKPNMSASCKEIVDSYLPVILDIKIGEMS 126

Qy 121 RPEVCSALNLCESLQKHLAELNHQKLESNKIPELDMTEVAVPFMANIPLLLYPQDGP 180  
Db 127 RPEVCSALNLCESLQKHLAELNHQKLESNKIPELDMTEVAVPFMANIPLLLYPQDGP 186

Qy 181 SKQPKDNGVCDQCIQWVTDIQTAVRTNSTFVQALVEHVKEECDRILGPGWADICKNYIS 240  
Db 187 SKQPKDNGVCDQCIQWVTDIQTAVRTNSTFVQALVEHVKEECDRILGPGWADICKNYIS 246

Qy 241 QYSEIATQMMHMM--OPKEICALVGFCDVEKEMPMOTLPAKVASKNVI PALELVEPIK 297  
Db 247 QYSEIATQMMHMMQDQPKKEICALVGFCDVEKEMPMOTLPAKVASKNVI PALELVEPIK 306

Qy 298 KHEVPAKSDVYCEVCEFLVKEVTKLIDNNKTEKILDAFDKMSKLPKSLSEECQEVVDT 357  
Db 307 KHEVPAKSDVYCEVCEFLVKEVTKLIDNNKTEKILDAFDKMSKLPKSLSEECQEVVDT 366

Qy 358 YGSSILSILLLEEVSPELVCSMLHLCSTGLPALTVHTVTPKDGGFCEVCKKLGYLDRL 417  
Db 367 YGSSILSILLLEEVSPELVCSMLHLCSTGLPALTVHTVTPKDGGFCEVCKKLGYLDRL 426

Qy 418 EKSTKOEILAALEKSGSFLPDYKQCDQFVAEYEPVLIILVEVMDPSFVCLKIGACP 477  
Db 427 EKSTKOEILAALEKSGSFLPDYKQCDQFVAEYEPVLIILVEVMDPSFVCLKIGACP 486

Qy 478 SAHKPLLGTEKICWGPSYWCQNTETAACNAVEHCKRHVWN 518  
Db 487 SAHKPLLGTEKICWGPSYWCQNTETAACNAVEHCKRHVWN 527

RESULT 6  
US-10-060-036-73  
; Sequence 73, Application US/10060036  
; Publication No. US20030073144A1  
; GENERAL INFORMATION:  
; APPLICANT: Benson, Darin R.  
; APPLICANT: Kalos, Michael D.  
; APPLICANT: Lodes, Michael J.  
; APPLICANT: Persing, David H.  
; APPLICANT: Hepler, William T.  
; APPLICANT: Jiang, Yugu  
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY  
; FILE REFERENCE: 210121.566  
; CURRENT APPLICATION NUMBER: US/10/060,036  
; CURRENT FILING DATE: 2002-01-30  
; NUMBER OF SEQ ID NOS: 4560  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 73  
; LENGTH: 527  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-060-036-73

Query Match 53.7%; Score 278; DB 14; Length 527;  
Best Local Similarity 99.4%; Pred. No. 1.9e-263;

Matches	518;	Conservative	0;	Mismatches	0;	Indels	3;	Gaps	1;																																											
QY	1	LASLLGAALAGPV	LGLKE	TRGSAYW	QNVKXTAS	DGGA	VKHC	LQTVW	NKPTVKSLP	CDIC 60																																										
Db	7	LASLLGAALAGPV	LGLKE	TRGSAYW	QNVKXTAS	DGGA	VKHC	LQTVW	NKPTVKSLP	CDIC 66																																										
QY	61	KDVVTTAAGDML	KDNATEE	EEILV	LEKTD	CDWL	P	KPNNS	ASCKE	IVDSYLPV	ILDI	IKGEMS 120																																								
Db	67	KDVVTTAAGDML	KDNATEE	EEILV	LEKTD	CDWL	P	KPNNS	ASCKE	IVDSYLPV	ILDI	IKGEMS 126																																								
QY	121	RGEVCSALN	CESLQ	KHLAE	LNH	QK	LES	NKIP	ELDM	TEV	VAF	PMANIP	LLLY	PQDGPR 180																																						
Db	127	RGEVCSALN	CESLQ	KHLAE	LNH	QK	LES	NKIP	ELDM	TEV	VAF	PMANIP	LLLY	PQDGPR 186																																						
QY	181	SKPQPKDNGD	VQCDCI	QMWTDI	QTA	VR	T	N	S	T	F	V	QAL	VEH	VKE	CDRL	GP	GMAD	ICK	KNYIS 240																																
Db	187	SKPQPKDNGD	VQCDCI	QMWTDI	QTA	VR	T	N	S	T	F	V	QAL	VEH	VKE	CDRL	GP	GMAD	ICK	KNYIS 246																																
QY	241	QYSEIAI	QMMH	M---	--	Q	P	K	E	I	C	A	L	V	G	F	D	E	V	K	E	N	P	M	O	T	L	V	P	A	K	V	A	S	K	N	V	I	P	A	L	E	L	V	E	P	I	K 297				
Db	247	QYSEIAI	QMMH	M	O	Q	P	K	E	I	C	A	L	V	G	F	D	E	V	K	E	N	P	M	O	T	L	V	P	A	K	V	A	S	K	N	V	I	P	A	L	E	L	V	E	P	I	K 306				
QY	298	KHEVPAKSD	V	N	V	C	E	V	C	E	F	L	V	K	E	T	K	I	L	D	N	N	K	T	E	K	I	L	D	A	F	D	K	M	C	S	K	L	P	K	S	L	S	E	C	Q	E	V	D	T 357		
Db	307	KHEVPAKSD	V	N	V	C	E	V	C	E	F	L	V	K	E	T	K	I	L	D	N	N	K	T	E	K	I	L	D	A	F	D	K	M	C	S	K	L	P	K	S	L	S	E	C	Q	E	V	D	T 366		
QY	358	YGSSILS	I	L	L	E	V	E	P	A	P	E	L	V	C	S	M	L	H	C	S	G	T	R	P	A	L	T	H	V	T	O	P	K	D	G	G	F	C	E	V	K	K	L	V	G	Y	L	D	R	N	L 417
Db	367	YGSSILS	I	L	L	E	V	E	P	A	P	E	L	V	C	S	M	L	H	C	S	G	T	R	P	A	L	T	H	V	T	O	P	K	D	G	G	F	C	E	V	K	K	L	V	G	Y	L	D	R	N	L 426
QY	418	EKNST	K	Q	B	E	I	L	A	E	K	G	S	F	L	P	D	P	Y	Q	K	O	C	D	O	F	V	A	E	Y	E	P	V	L	E	I	L	V	E	M	D	S	F	V	C	L	K	I	G	A	C	P 477
Db	427	EKNST	K	Q	B	E	I	L	A	E	K	G	S	F	L	P	D	P	Y	Q	K	O	C	D	O	F	V	A	E	Y	E	P	V	L	E	I	L	V	E	M	D	S	F	V	C	L	K	I	G	A	C	P 486
QY	478	SAHPE	L	L	G	T	E	K	I	G	P	S	I	W	C	O	N	T	E	T	A	A	Q	C	N	A	V	E	H	C	K	R	H	V	N	518																
Db	487	SAHPE	L	L	G	T	E	K	I	G	P	S	I	W	C	O	N	T	E	T	A	A	Q	C	N	A	V	E	H	C	K	R	H	V	N	527																

RESULT 7  
US-09-978-418-40  
; Sequence 40, Application US/09978418  
; Publication No. US2003011897A1  
; GENERAL INFORMATION:  
; APPLICANT: Bejanin, Stephan  
; APPLICANT: Tanaka, Hiroaki  
; TITLE OF INVENTION: HUMAN CDNAS AND PROTEINS AND USES THEREOF  
; FILE REFERENCE: 142.US5.REG  
; CURRENT APPLICATION NUMBER: US/09/978,418  
; CURRENT FILING DATE: 2002-10-15  
; PRIOR APPLICATION NUMBER: 60/311,305  
; PRIOR FILING DATE: 2001-08-10  
; PRIOR APPLICATION NUMBER: 60/314,734  
; PRIOR FILING DATE: 2001-08-24  
; PRIOR APPLICATION NUMBER: 60/318,204  
; PRIOR FILING DATE: 2001-09-07  
; PRIOR APPLICATION NUMBER: 60/326,470  
; PRIOR FILING DATE: 2001-10-01  
; NUMBER OF SEQ ID NOS: 52  
; SOFTWARE: JPatent  
; SEQ ID NO 40  
; LENGTH: 479  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-978-418-40

	Query Match	49.0%;	Score 254;	DB 10;	Length 479;	
	Best Local Similarity	100.0%;	Pred. No. 5.7e-240;			
	Matches 254;	Conservative	0;	Mismatches	0;	Gaps 0;
Qy	1	LASLILGAALAGFVLGLKECTRGSAVWCONVKTSAPCGAVKICLOTQVWNKPVTXSLPCDIC				60
db	7	LASLILGAALAGFVLGLKECTRGSAVWCONVKTSAPCGAVKICLOTQVWNKPVTXSLPCDIC				66

```
; PRIOR APPLICATION NUMBER: 08/958,970
; PRIOR FILING DATE: 1997-10-28
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 523
; TYPE: PRT
; ORGANISM: homo sapiens
US-09-767-007A-2

Query Match      36.1%; Score 187; DB 9; Length 523;
Best Local Similarity 100.0%; Pred. No. 2.9e-174;
Matches 187; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 LASLGAALAGPVLGLKECTRGSAVWQNVKTSASDCGAVKHCLQTVWKNKPTVKSLPCDIC 60
Db 7 LASLGAALAGPVLGLKECTRGSAVWQNVKTSASDCGAVKHCLQTVWKNKPTVKSLPCDIC 66

Qy 61 KDVTAAAGDMLKDNATBEELVYLEKTCDWLPKPNMGASCKEIVDSYLPVILDIKIGEMS 120
Db 67 KDVTAAAGDMLKDNATBEELVYLEKTCDWLPKPNMGASCKEIVDSYLPVILDIKIGEMS 126

Qy 121 RPGEVCSALNLCESLQKHLAELNHOKLESNKIPELDMTEVYVAPFMANIPLILYPODGPR 180
Db 127 RPGEVCSALNLCESLQKHLAELNHOKLESNKIPELDMTEVYVAPFMANIPLILYPODGPR 186

Qy 181 SKPOPKD 187
Db 187 SKPOPKD 193

RESULT 10
US-09-767-007A-3
; Sequence 3, Application US/09767007A
; Patent No. US2002007275A1
; GENERAL INFORMATION:
; APPLICANT: John S. O'Brien
; APPLICANT: Yasuo Kishimoto
; TITLE OF INVENTION: PHARMACEUTICAL COMPOSITIONS COMPRISING
; TITLE OF INVENTION: SAPOSIN C AND NEUROTROPIC PEPTIDES DERIVED THEREFROM
; FILE REFERENCE: MYELOS.2DC1C1
; CURRENT APPLICATION NUMBER: US/09/767,007A
; CURRENT FILING DATE: 2001-01-22
; PRIOR APPLICATION NUMBER: 08/958,970
; PRIOR FILING DATE: 1997-10-28
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 80
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Artificial peptide
US-09-767-007A-3

Query Match      15.4%; Score 80; DB 9; Length 80;
Best Local Similarity 100.0%; Pred. No. 4e-70;
Matches 80; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 305 SDVYCEVCEFLVKEVTKLIDNNKTEKILDAFDKMCCLKPKSLSEECQEVVDYTGSSILS 364
Db 1 SDVYCEVCEFLVKEVTKLIDNNKTEKILDAFDKMCCLKPKSLSEECQEVVDYTGSSILS 60

Qy 365 ILLEEVSPELVCSMLHLCSG 384
Db 61 ILLEEVSPELVCSMLHLCSG 80

RESULT 11
US-09-753-126-3
; Sequence 3, Application US/09753126
; Patent No. US2002012719A1
; GENERAL INFORMATION:
; APPLICANT: OKKELS, JENS SIGURD
; APPLICANT: JENSEN, ANNE DAM
; APPLICANT: HALKIER, TORBEN
; APPLICANT: JENSEN, RIKKE BOLDING
; TITLE OF INVENTION: IMPROVED LYSOSOMAL ENZYMES AND LYSOSOMAL ENZYME
; TITLE OF INVENTION: ACTIVATORS
; FILE REFERENCE: 31-0006000S
; CURRENT APPLICATION NUMBER: US/10/330,697
; CURRENT FILING DATE: 2002-12-27
; PRIOR APPLICATION NUMBER: US/09/753,126
; PRIOR FILING DATE: 2001-06-11
; PRIOR APPLICATION NUMBER: PA 1999 01891
; PRIOR FILING DATE: 1999-12-30
; PRIOR APPLICATION NUMBER: 60/174,652
; PRIOR FILING DATE: 2000-01-06
; PRIOR APPLICATION NUMBER: PA 200 00865
; PRIOR FILING DATE: 2000-06-02
; PRIOR APPLICATION NUMBER: 60/210,984
; PRIOR FILING DATE: 2000-06-12
; PRIOR APPLICATION NUMBER: 60/211,124
; PRIOR APPLICATION NUMBER: 2000-06-12
; PRIOR APPLICATION NUMBER: PA 2000 01027
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: 60/217,497
```





; ORGANISM: Homo sapiens  
US-09-780-438A-1

Query Match 7.7%; Score 40; DB 10; Length 40;  
Best Local Similarity 100.0%; Pred. No. 3.4e-31;  
Matches 40; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 305 SDVYCEVCEFLVKEVTKLIDNNKTEKEILDADFDMCKSLP 344  
|||||  
Db 1 SDVYCEVCEFLVKEVTKLIDNNKTEKEILDADFDMCKSLP 40  
|||||

Search completed: May 5, 2004, 13:43:38  
Job time : 41.7582 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 5, 2004, 13:42:09 ; Search time 16.405 Seconds  
(without alignments)  
3037.320 Million cell updates/sec

Title: US-09-743-684A-1\_COPY\_7\_524

Perfect score: 518

Sequence: 1 LASLLGALAGPVLGLKCT.....NTETAQCNAVEHCKRHVMN 518

Scoring table: OLIGO  
Gapop 60.0 , Gapext 60.0

Searched: 283366 seqs, 96191526 residues

Word size : 0

Total number of hits satisfying chosen parameters: 283366

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : PIR\_78.\*

- 1: pir1.\*
- 2: pir2.\*
- 3: pir3.\*
- 4: pir4.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	278	53.7	527	1 SAHUP	saposin precursor
2	22	4.2	79	2 A49475	cerebroside sulfat
3	22	4.2	554	1 A28716	saposin precursor
4	22	4.2	557	1 JH0604	saposin precursor
5	13	2.5	80	2 S21770	saposin-C - bovine
6	9	1.7	81	2 A32026	glucosylceramide b
7	8	1.5	177	2 T47133	hypothetical prote
8	8	1.5	186	1 HHFF23	heat shock protein
9	8	1.5	186	2 B20647	heat shock protein
10	8	1.5	201	2 S36451	fimbrial protein f
11	8	1.5	201	2 S01929	fimbrial protein p
12	8	1.5	216	2 F64390	coenzyme F420 hydr
13	8	1.5	326	2 B90841	probable ABC trans
14	8	1.5	326	2 A85699	probable ABC trans
15	8	1.5	354	2 A23614	probable undecapre
16	8	1.5	435	2 D69992	ABC transporter (m
17	8	1.5	655	2 D85988	hypothetical prote
18	8	1.5	655	2 B65116	hypothetical 73.6
19	8	1.5	655	2 A91143	hypothetical prote
20	8	1.5	813	1 A49123	fibroblast growth
21	8	1.4	62	2 S77874	probable preprotei
22	7	1.4	94	2 A89986	hypothetical prote
23	7	1.4	105	2 S46149	ribosomal protein
24	7	1.4	112	2 A37203	lens fiber membran
25	7	1.4	125	2 B83857	chorismate mutase
26	7	1.4	148	2 JC2111	helix-loop-helix p
27	7	1.4	154	2 S76225	hypothetical prote
28	7	1.4	158	2 E84956	acetylacate synth
29	7	1.4	161	2 H84289	transcription regu

RESULT 1

SAHUP

Saposin precursor [validated] - human

N;Alternate names: cerebroside sulfate activator protein; co-beta-glucosidase; component ein (SAP); sphingolipid activator protein A2; sulfate sulfatase activator protein N;Contains: prosaposin; saposin A; saposin B; saposin C; saposin D

C;Species: Homo sapiens (man)

C;Date: 30-Jun-1992 #sequence revision 17-Nov-1995 #text change 08-Dec-2000

C;Accession: JX0061; A57368; A42003; B42003; C42003; D42003; A30367; S34740; S36140; S360226; I37265; I37264

R;Nakano, T.; Sandhoff, K.; Stuenkel, J.; Christomanou, H.; Suzuki, K.

J. Biochem. 105, 152-154, 1989

A;Title: Structure of full-length cDNA coding for sulfate activator, a Co-beta-glucosidase

A;Reference number: JX0061; MUID:89255151; PMID:2498298

A;Accession: JX0061

A;Molecule type: mRNA

A;Residues: 1-527 <NAK>

A;Cross-references: GB:D00422; NID:G220063; PIDN:BAA00321.1; PID:G220064

A;Note: alternative splice form 1

A;Accession: A57368

A;Molecule type: mRNA

A;Residues: 1-259,263-527 <NA2>

A;Cross-references: GB:J03015; GB:J03086; NID:G337755; PIDN:AAB59494.1; PID:G337756

A;Note: alternative splice form 2

R;Rorman, E.G.; Scheinker, V.; Grabowski, G.A.

Genomics 13, 312-318, 1992

A;Title: Structure and evolution of the human prosaposin chromosomal gene.

A;Reference number: A42003; MUID:92307663; PMID:1612590

A;Accession: A42003

A;Molecule type: DNA

A;Residues: 50-140 <ROR>

A;Cross-references: GB:M86181

A;Note: sequence extracted from NCBI backbone (NCBIN:107235, NCBIP:107236)

A;Accession: B42003

A;Molecule type: DNA

A;Residues: 185-259;263-276 <RO2>

A;Note: sequence extracted from NCBI backbone (NCBIN:107235, NCBIP:107237)

A;Accession: C42003

A;Molecule type: DNA

A;Residues: 305-393 <RO3>

A;Note: sequence extracted from NCBI backbone (NCBIN:107235, NCBIP:107238); sequence inc.

A;Accession: D42003

A;Molecule type: DNA

A;Residues: 399-487 <RO4>

A;Note: sequence extracted from NCBI backbone (NCBIN:107235, NCBIP:107239); sequence inc.

R;Rorman, E.G.; Grabowski, G.A.

Genomics 5, 486-492, 1989

A;Title: Molecular cloning of a human co-beta-glucosidase cDNA: evidence that four sphin

A;Reference number: A30367; MUID:90129043; PMID:2515150

A;Accession: A30367

A;Molecule type: mRNA

A;Residues: 1-259,263-527 <RO5>

A;Cross-references: GB:J03077; NID:G183230; PIDN:AAA52560.1; PID:G183231  
 A;Note: alternative splice form 2  
 R;Hiraawa, M.; O'Brien, J.S.; Kishimoto, Y.; Galdickia, M.; Fluharty, A.L.; Ginns, E.I.; Arch. Biochem. Biophys. 304, 110-116, 1993  
 A;Title: Isolation, characterization, and proteolysis of human prosaposin, the precursor  
 A;Reference number: S34740; MUID:93311991; PMID:8323276  
 A;Accession: S34740  
 A;Molecule type: protein  
 A;Residues: 17-24;165-172;180-189;301-305 <HIR>  
 R;Tyynelae, J.; Palmer, D.N.; Baumann, M.; Haltia, M. FEBS Lett. 330, 8-12, 1993  
 A;Title: Storage of saposins A and D in infantile neuronal ceroid-lipofuscinosis.  
 A;Reference number: S36140; MUID:93380576; PMID:8370464  
 A;Accession: S36140  
 A;Molecule type: protein  
 A;Residues: 'XX', 62, 'X', 64-65, 'X', 67-79, 'X', 81-84 <TY>  
 A;Note: saposin A  
 A;Accession: S36141  
 A;Molecule type: protein  
 A;Residues: 'XXX', 413-414, 'X', 416-428, 'X', 430-434 <TY2>  
 A;Note: saposin D  
 R;Holtzman, H.; Sandhoff, K.; Kwon, H.Y.; Harzer, K.; Nakano, T.; Suzuki, K. J. Biol. Chem. 266, 7556-7560, 1991  
 A;Title: Sulfatide activator protein. Alternative splicing that generates three mRNAs and  
 A;Reference number: S36988; MUID:91210267; PMID:2019586  
 A;Accession: S36988  
 A;Status: nucleic acid sequence not shown; translation not shown  
 A;Molecule type: mRNA  
 A;Residues: 1-240, 'S', 242-527 <HOL>  
 A;Cross-references: EMBL:M60255; NID:G337759; PIDN:AAA36594.1; PID:G337760  
 A;Note: the nucleotide sequence was submitted to the EMBL Data Library, June 1991  
 A;Note: cerebroside sulfate activator protein mutant MU-9; corresponds to alternative splicing  
 A;Accession: S36989  
 A;Status: nucleic acid sequence not shown; translation not shown  
 A;Molecule type: mRNA  
 A;Residues: 1-240, 'S', 242-259, 263-527 <HO2>  
 A;Cross-references: EMBL:M60257; NID:G337764; PIDN:AAA36595.1; PID:G337765  
 A;Note: the nucleotide sequence was submitted to the EMBL Data Library, June 1991  
 A;Note: cerebroside sulfate activator protein mutant MU-0; corresponds to alternative splicing  
 A;Accession: S36990  
 A;Status: nucleic acid sequence not shown; translation not shown  
 A;Molecule type: mRNA  
 A;Residues: 1-240, 'S', 242-259, 261-527 <HO3>  
 A;Cross-references: EMBL:M60258; NID:G337766; PIDN:AAA36596.1; PID:G337767  
 A;Note: the nucleotide sequence was submitted to the EMBL Data Library, June 1991  
 A;Note: cerebroside sulfate activator protein mutant MU-6; corresponds to alternative splicing  
 R;Kondoh, K.; Hinenio, T.; Sano, A.; Kakimoto, Y. Blochem. Biophys. Res. Commun. 181, 286-292, 1991  
 A;Title: Isolation and characterization of prosaposin from human milk.  
 A;Reference number: PS0330; MUID:92068206; PMID:1958198  
 A;Accession: PS0330  
 A;Molecule type: protein  
 A;Residues: 17-24, 'X', 26 <KON>  
 A;Experimental source: milk  
 R;Kretz, K.A.; Carson, G.S.; Morimoto, S.; Kishimoto, Y.; Fluharty, A.L.; O'Brien, J.S. Proc. Natl. Acad. Sci. U.S.A. 87, 2541-2544, 1990  
 A;Title: Characterization of a mutation in a family with saposin B deficiency: a glycosyl  
 A;Reference number: A35985; MUID:90207231; PMID:2320574  
 A;Accession: A35985  
 A;Molecule type: mRNA  
 A;Residues: 213-221 <KRE>  
 A;Cross-references: GB:M32221  
 A;Accession: B35985  
 A;Status: nucleic acid sequence not shown; translation not shown  
 A;Molecule type: mRNA  
 A;Residues: 1-259, 263-527 <KR2>  
 A;Cross-references: GB:N32221; NID:G337761; PIDN:AAA60303.1; PID:G337762  
 A;Experimental source: lymphoblast  
 A;Accession: C35985  
 A;Molecule type: mRNA  
 A;Residues: 213-216, 'I', 218-221 <KR3>  
 A;Note: sequence from patients with activator-deficient metachromatic leukodystrophy; th  
 R;Fuerst, W.; Schubert, J.; Machleidt, W.; Meyer, H.E.; Sandhoff, K.

Eur. J. Biochem. 192, 709-714, 1990  
 A;Title: The complete amino-acid sequences of human ganglioside GM2 activator protein and  
 A;Reference number: S13195; MUID:9106165; PMID:2209618  
 A;Accession: S13196  
 A;Molecule type: protein  
 A;Residues: 195-259, 263-277 <FUE>  
 R;Morimoto, S.; Martin, B.M.; Yamamoto, Y.; Kretz, K.A.; O'Brien, J.S.; Kishimoto, Y. Proc. Natl. Acad. Sci. U.S.A. 86, 3389-3393, 1989  
 A;Title: Saposin A: second cerebroside activator protein.  
 A;Reference number: A32784; MUID:89240739; PMID:2717620  
 A;Accession: A32784  
 A;Molecule type: protein  
 A;Residues: 60-84;86-107;109-119;125-134 <MOR>  
 R;O'Brien, J.S.; Kretz, K.A.; Dewji, N.; Wenger, D.A.; Esch, F.; Fluharty, A.L. Science 241, 1098-1101, 1988  
 A;Title: Coding of two sphingolipid activator proteins (SAP-1 and SAP-2) by same genetic  
 A;Reference number: A41240; MUID:88321660; PMID:2842863  
 A;Accession: A41240  
 A;Molecule type: mRNA  
 A;Residues: 'GSSR', 18-259, 263-299, 'D', 301-302, 'D', 304-527 <OAB>  
 A;Cross-references: GB:J03086  
 R;Dewji, N.N.; Wenger, D.A.; O'Brien, J.S. Proc. Natl. Acad. Sci. U.S.A. 84, 8552-8556, 1987  
 A;Title: Nucleotide sequence of cloned cDNA for human sphingolipid activator protein 1 p  
 A;Reference number: S02289; MUID:88068647; PMID:2825202  
 A;Accession: S02289  
 A;Status: significant sequence differences  
 A;Molecule type: mRNA  
 A;Cross-references: EMBL:J03015  
 A;Note: this sequence corrected by A41240  
 R;Kleinschmidt, T.; Christomanou, H.; Braunitzer, G. Biol. Chem. Hoppe-Seyler 369, 1361-1365, 1988  
 A;Title: Complete amino-acid sequence of the naturally occurring A(2) activator protein f  
 A;Reference number: S02028; MUID:89207118; PMID:3242555  
 A;Accession: S02028  
 A;Molecule type: protein  
 A;Residues: 195-259, 263-276 <KLB>  
 R;Fuerst, W.; Machleidt, W.; Sandhoff, K. Biol. Chem. Hoppe-Seyler 369, 317-328, 1988  
 A;Title: The precursor of sulfatide activator protein is processed to three different pr  
 A;Reference number: S00813; MUID:89000190; PMID:3048308  
 A;Accession: S00813  
 A;Molecule type: protein  
 A;Residues: 410-487 <FU2>  
 R;Kleinschmidt, T.; Christomanou, H.; Braunitzer, G. Biol. Chem. Hoppe-Seyler 368, 1571-1578, 1987  
 A;Title: Complete amino-acid sequence and carbohydrate content of the naturally occurring  
 A;Reference number: S00226; MUID:88163077; PMID:3442600  
 A;Accession: S00226  
 A;Molecule type: protein  
 A;Residues: 314-333 <KL2>  
 R;Vaccaro, A.M.; Salvio, R.; Barca, A.; Tatti, M.; Ciaffoni, F.; Maras, B.; Siciliano, J. Biol. Chem. 270, 9953-9960, 1995  
 A;Title: Structural analysis of saposin C and B. Complete localization of disulfide bridg  
 A;Reference number: A57297; MUID:95247790; PMID:7730378  
 A;Contents: annotation; disulfide bonds; glycosylation  
 R;Holtzman, H.; Sandhoff, K.; Fuerst, W.; Kwon, H.Y.; Schnabel, D.; Suzuki, K. FEBS Lett. 280, 267-270, 1991  
 A;Title: The organization of the gene for the human cerebroside sulfate activator protein  
 A;Reference number: I37264; MUID:91192146; PMID:2013321  
 A;Accession: I37265  
 A;Status: translated from GB/EMBL/DBJ  
 A;Molecule type: DNA  
 A;Residues: 59-125 <RES>  
 A;Cross-references: EMBL:X57107; NID:G30234; PIDN:CAA40391.1; PID:G30235  
 A;Accession: I37264  
 A;Status: translated from GB/EMBL/DBJ  
 A;Molecule type: DNA  
 A;Residues: 307-516 <RE2>  
 A;Cross-references: EMBL:X57108; NID:G30232; PIDN:CAA40392.1; PID:G1565257  
 A;Note: sequence revised relative to PID:G30233 (corrected coding region)  
 C;Genetics:

A:Gene: GDB:PSAP; GLBA  
A:Cross-references: GDB:120366; OMIM:176801  
A:Map position: 10q22.1-10q22.1  
A:Introns: 83/3; 338/3; 401/1; 453/3; 480/3  
A>Note: defects in this gene may cause variant Gaucher disease, variant Tay-Sachs disease  
A>Note: list of introns is incomplete

Query Match 53.7%; Score 278; DB 1; Length 527;  
Best Local Similarity 99.4%; Pred. No. 1.1e-281;  
Matches 518; Conservative 0; Mismatches 0; Indels 3; Gaps 1;

Qy 1 LASLGAALAGPVLGLKECTRGSAWCONVTASDCGAVKHCLQTVWKNKPTVKSIPCDIC 60  
Db 7 LASLGAALAGPVLGLKECTRGSAWCONVTASDCGAVKHCLQTVWKNKPTVKSIPCDIC 66  
Qy 61 KDVVTAAGDMLKDNATEEILVLEKTCMDLVPKPNMSASCKEIVDSYLPVILDIKIGEMS 120  
Db 67 KDVVTAAGDMLKDNATEEILVLEKTCMDLVPKPNMSASCKEIVDSYLPVILDIKIGEMS 126  
Qy 121 RPGEVCSALNLCESLQKHLAELNHQKLESNKIPELDMTEVVAFPMANIPILLYPQDGPR 180  
Db 127 RPGEVCSALNLCESLQKHLAELNHQKLESNKIPELDMTEVVAFPMANIPILLYPQDGPR 186  
Qy 181 SKPOPKNQDGYCDQCIQWVTDIQAVRTNSTFVQALVEHVKECDRLGPGWADICKNYIS 240  
Db 187 SKPOPKNQDGYCDQCIQWVTDIQAVRTNSTFVQALVEHVKECDRLGPGWADICKNYIS 246  
Qy 241 QYSEIAIQMMHMM---OPKEICALVGFCDVEKEMPMQTLVPAKVASKNVIPALELVEPIK 297  
Db 247 QYSEIAIQMMHMQDQPKKEICALVGFCDVEKEMPMQTLVPAKVASKNVIPALELVEPIK 306  
Qy 298 KHEVPAKSDVYCEVCEFLVKEVTKLIDNNKTEKEILDADFDMCKSLPKSLSEECQEVVDT 357  
Db 307 KHEVPAKSDVYCEVCEFLVKEVTKLIDNNKTEKEILDADFDMCKSLPKSLSEECQEVVDT 366  
Qy 358 YGSSILSTLLBEVSPVLVCSHLHLCGTRLPALTAVHTVTPKDGCFCEVCKLVGLDRNL 417  
Db 367 YGSSILSTLLBEVSPVLVCSHLHLCGTRLPALTAVHTVTPKDGCFCEVCKLVGLDRNL 426  
Qy 418 EKNSTKOEILAALEKGCSEFLPDYQKQCDQFVAEYEPVLIILVEVMDPSPFVCLIKIGACP 477  
Db 427 EKNSTKOEILAALEKGCSEFLPDYQKQCDQFVAEYEPVLIILVEVMDPSPFVCLIKIGACP 486  
Qy 478 SAHKPLLGTEKICNGPSYWCQNTETAACQNAVEHCKRHVWN 518  
Db 487 SAHKPLLGTEKICNGPSYWCQNTETAACQNAVEHCKRHVWN 527

RESULT 2  
A49475  
cerebroside sulfate activator - pig  
C:Species: Sus scrofa domestica (domestic pig)  
C:Date: 24-Feb-1994 #sequence\_revision 18-Nov-1994 #text\_change 17-May-1996  
C:Accession: A49475  
R:Stevens, R.L.; Faull, K.F.; Conklin, K.A.; Green, B.N.; Fluharty, A.L.  
Biochemistry 32, 4051-4059, 1993  
A:Title: Porcine cerebroside sulfate activator: further structural characterization and  
A:Reference number: A49475; MUID:93229506; PMID:8471613  
A:Accession: A49475  
A:Status: preliminary  
A:Molecule type: protein  
A:Residues: 1-79 <STE>  
A:Experimental source: kidney  
A>Note: sequence extracted from NCBI backbone (NCBIP:129597)  
C:Superfamily: saposin; saposin repeat homology  
F:1-79/Domain: saposin repeat homology <SAP>

Query Match 4.2%; Score 22; DB 2; Length 79;  
Best Local Similarity 100.0%; Pred. No. 5.1e-15;  
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 235 CKNVISQYSEIAIQMMHMQPK 256  
Db 235 CKNVISQYSEIAIQMMHMQPK 256

Db 47 CKNVISQYSEIAIQMMHMQPK 68

RESULT 3  
A28716  
saposin precursor - rat  
N:Alternate names: cerebroside sulfate activator protein; co-beta-glucosidase; component  
ein (SAP); sphingolipid activator protein A2; sulfated glycoprotein 1; sulfatide sulfata  
N:Contains: prosaposin; saposin A; saposin B; saposin C; saposin D  
C:Species: Rattus norvegicus (Norway rat)  
C:Date: 10-Sep-1999 #sequence\_revision 10-Sep-1999 #text\_change 10-Sep-1999  
C:Accession: A28716  
R:Collard, M.W.; Sylvester, S.R.; Tsuruta, J.K.; Griswold, M.D.  
Biochemistry 27, 4557-4564, 1988  
A:Title: Biosynthesis and molecular cloning of sulfated glycoprotein 1 secreted by rat  
A:Reference number: A28716; MUID:89000647; PMID:3048385  
A:Accession: A28716  
A:Molecule type: mRNA  
A:Residues: 1-554 <COL>  
A:Cross-references: GB:M19936; NID:g206904; PIDN:AAA42136.1; PID:g206905  
A>Note: parts of this sequence, including the amino end of the mature protein, were dete  
C:Function:  
A:Description: saposins bind sphingolipids, form hydrophilic complexes and make them acc  
A:Pathway: sphingolipid catabolism  
A>Note: saposins A and C (SAP-2) activate hydrolysis of glucocerebroside by beta-glucosy  
A>Note: saposin B (SAP-1) activates hydrolysis of galactocerebroside sulfate by arylsulfi  
A>Note: saposin D activates hydrolysis of sphingomyelin by sphingomyelin phosphodiesterase  
C:Superfamily: saposin; saposin repeat homology  
C:Keywords: alternative splicing; glycoprotein; lysosomal storage disease; lysosome; sph  
F:1-16/Domain: signal sequence #status predicted <SIG>  
F:17-554/Product: prosaposin #status predicted <PRO>  
F:55-148/Domain: saposin repeat homology <SAP1>  
F:60-143/Product: saposin A #status predicted <SAPA>  
F:189-280/Domain: saposin repeat homology <SAP2>  
F:194-273/Product: saposin B #status predicted <SAB1>  
F:306-397/Domain: saposin repeat homology <SAP3>  
F:310-397/Product: saposin C #status predicted <SAPC>  
F:431-522/Domain: saposin repeat homology <SAP4>  
F:437-514/Product: saposin D #status predicted <SAPD>  
F:63-138,66-132,94-106,439-512,442-506,470-481/Disulfide bonds: #status predicted  
F:80,214,331,456/Binding site: carbohydrate (Asn) (covalent) #status predicted  
F:197-270,200-264,228-240,314-387,317-381,345-356/Disulfide bonds: #status predicted

Query Match 4.2%; Score 22; DB 1; Length 554;  
Best Local Similarity 100.0%; Pred. No. 3.2e-14;  
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 425 EILAALEKGCSEFLPDYQKQCD 446  
Db 461 EILAALEKGCSEFLPDYQKQCD 482

RESULT 4  
JH0604  
saposin precursor - mouse  
N:Alternate names: cerebroside sulfate activator protein; co-beta-glucosidase; component  
ein (SAP); sphingolipid activator protein A2; sulfated glycoprotein 1; sulfatide sulfata  
N:Contains: prosaposin; saposin A; saposin B; saposin C; saposin D  
C:Species: Mus musculus (house mouse)  
C:Date: 10-Sep-1999 #sequence\_revision 10-Sep-1999 #text\_change 10-Sep-1999  
C:Accession: JH0604  
R:Tsuda, M.; Sakiyama, T.; Endo, H.; Kitagawa, T.  
Biochem. Biophys. Res. Commun. 184, 1266-1272, 1992  
A:Title: The primary structure of mouse saposin.  
A:Reference number: JH0604; MUID:92272718; PMID:1590798  
A:Accession: JH0604  
A:Molecule type: mRNA  
A:Residues: 1-557 <TSU>  
A:Cross-references: GB:S36200; NID:g249386; PIDN:AAB22175.1; PID:g249387  
A:Experimental source: liver  
C:Function:  
A:Description: saposins bind sphingolipids, form hydrophilic complexes and make them acc  
A:Pathway: sphingolipid catabolism

A>Note: saposins A and C (SAP-2) activate hydrolysis of glucocerebroside by beta-glucosyl  
A>Note: saposin B (SAP-1) activates hydrolysis of galactocerebroside sulfate by arylsulfa  
C:Superfamily: saposin; saposin hydrolysis of sphingomyelin by sphingomyelin phosphodiesterase  
C:Keywords: alternative splicing; glycoprotein; lysosomal storage disease; lysosome; sph  
F:1-16/Domain: signal sequence #status predicted <SIG>  
F:17-55/Product: prosaposin #status predicted <PRO>  
F:55-148/Domain: prosaposin #status predicted <SIG>  
F:160-143/Domain: prosaposin #status predicted <PRO>  
F:189-283/Domain: saposin A #status predicted <SAPA>  
F:194-276/Product: saposin A repeat homology <SAP>  
F:309-400/Domain: saposin B #status predicted <SAB1>  
F:333-392/Domain: saposin C #status predicted <SAPC>  
F:434-525/Domain: saposin repeat homology <SAP4>  
F:440-517/Product: saposin D #status predicted <SAPD>  
F:63-138,66-132,94-106,197-273,200-267,229-240,317-330,320-384,348-359,442-515,445-509,4  
F:80,214,334,379,459/Binding site: carbonylate (Asn) (covalent) #status predicted

Query Match 4.2%; Score 22; DB 1; Length 557;  
Best Local Similarity 100.0%; Pred. No. 3.2e-14;  
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 425 EILAALEKGCSELPDPYKQCD 446  
Db 464 EILAALEKGCSELPDPYKQCD 485

RESULT 5  
S21770  
saposin-C - bovine  
C:Species: Bos primigenius taurus (cattle)  
C:Date: 19-Mar-1997 #sequence\_revision 19-Mar-1997 #text\_change 31-Oct-1997  
C:Accession: S21770  
R:Sano, A.; Mizuno, T.; Kondoh, K.; Hinenio, T.; Ueno, S.; Kakimoto, Y.; Morita, N.  
Biochim. Biophys. Acta 1120, 75-80, 1992  
A:Title: Saposin-C from bovine spleen: complete amino acid sequence and relation between  
A:Reference number: S21770; MUID:92207994; PMID:1554743  
A:Accession: S21770  
A>Status: preliminary  
A:Molecule type: protein  
A:Residues: 1-80 <SAN>  
C:Superfamily: saposin; saposin repeat homology

Query Match 2.5%; Score 13; DB 2; Length 80;  
Best Local Similarity 100.0%; Pred. No. 1.3e-05;  
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 371 SPELVCSMLHCS 383  
Db 67 SPELVCSMLHCS 79

RESULT 6  
A32026  
glucosylceramide beta-glucosidase activator protein SAP-2 - guinea pig  
C:Species: Cavia porcellus (guinea pig)  
C:Date: 07-Jun-1990 #sequence\_revision 07-Jun-1990 #text\_change 17-May-1996  
C:Accession: A32026  
R:Sano, A.; Radin, N.S.; Johnson, L.L.; Tarr, G.E.  
J. Biol. Chem. 263, 19597-19601, 1988  
A:Title: The activator protein for glucosylceramide beta-glucosidase from guinea pig liv  
A:Reference number: A32026; MUID:89066787; PMID:3198642  
A:Accession: A32026  
A:Molecule type: protein  
A:Residues: 1-81 <SAN>  
C:Superfamily: saposin; saposin repeat homology  
F:1-81/Domain: saposin repeat homology <SAP>

Query Match 1.7%; Score 9; DB 2; Length 81;  
Best Local Similarity 100.0%; Pred. No. 0.2;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 351 CQEVWDYTG 359

Db 47 CQEVWDYTG 55

RESULT 7  
T47133  
hypothetical protein DKFZp761P19121.1 - human (fragment)  
C:Species: Homo sapiens (man)  
C:Date: 20-Apr-2000 #sequence\_revision 20-Apr-2000 #text\_change 20-Apr-2000  
C:Accession: T47133  
R:Pousetka, A.; Wellenreuther, R.; Mewes, H.W.; Weil, B.; Wiemann, S.  
submitted to the Protein Sequence Database, March 2000  
A:Reference number: Z24374  
A:Accession: T47133  
A>Status: preliminary  
A:Molecule type: mRNA  
A:Residues: 1-177 <AAA>  
A:Cross-references: EMBL:AL162013  
A:Experimental source: adult amygdala; clone DKFZp761P19121  
C:Genetics:  
A>Note: DKFZp761P19121.1

Query Match 1.5%; Score 8; DB 2; Length 177;  
Best Local Similarity 100.0%; Pred. No. 4.6;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 425 EILAALEK 432  
Db 146 EILAALEK 153

RESULT 8  
HHPF23  
heat shock protein 23 - fruit fly (Drosophila melanogaster)  
C:Species: Drosophila melanogaster  
C:Date: 15-Nov-1984 #sequence\_revision 15-Nov-1984 #text\_change 16-Feb-1997  
C:Accession: A02919  
R:Ingolia, T.D.; Craig, E.A.  
Proc. Natl. Acad. Sci. U.S.A. 79, 2360-2364, 1982  
A:Title: Four small Drosophila heat shock proteins are related to each other and to mamme  
A:Reference number: A93909; MUID:82248004; PMID:6285380  
A:Accession: A02919  
A:Molecule type: DNA  
A:Residues: 1-186 <ING>  
A>Note: the authors translated the codon CCC for residue 25 as Leu, CGA for residue 32 as  
A>Note: the codon usage table proposed by the authors and the translation of the protein  
C:Comment: This small heat shock protein is related to alpha crystallin.  
C:Genetics:  
A:Gene: FlyBase:Hsp23  
A:Cross-references: FlyBase:FBgn0001224  
A:Map position: 3L (67B)  
C:Superfamily: alpha-crystallin  
C:Keywords: heat shock; stress-induced protein

Query Match 1.5%; Score 8; DB 1; Length 186;  
Best Local Similarity 100.0%; Pred. No. 4.8;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 166 MANIPLLL 173  
Db 1 MANIPLLL 8

RESULT 9  
B20647  
heat shock protein 23 - fruit fly (Drosophila melanogaster)  
C:Species: Drosophila melanogaster  
C:Date: 09-Sep-1987 #sequence\_revision 09-Sep-1987 #text\_change 21-Jul-2000  
C:Accession: B20647  
R:Southgate, R.; Ayme, A.; Voellmy, R.  
J. Mol. Biol. 165, 35-57, 1983  
A:Title: Nucleotide sequence analysis of the Drosophila small heat shock gene cluster at  
A:Reference number: A92897; MUID:83189140; PMID:6302284

A:Accession: B20647  
 A:Molecule type: DNA  
 A:Residues: 1-186 <SOU>  
 A:Cross-references: GB:V00210; NID:g8099; PIDN:CAA23494.1; PID:g8100  
 C:Genetics:  
 A:Gene: FlyBase:Hsp23  
 A:Cross-references: FlyBase:FBgn0001224  
 C:Superfamily: alpha-crystallin

Query Match 1.5%; Score 8; DB 2; Length 186;  
 Best Local Similarity 100.0%; Pred. No. 4.8;  
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 166 MANIPLLL 173  
 |||||  
 Db 1 MANIPLLL 8

## RESULT 10

S36451  
 fimbrial protein fimX - Bordetella bronchiseptica  
 C:Species: Bordetella bronchiseptica  
 C:Date: 06-Jan-1995 #sequence\_revision 06-Jan-1995 #text\_change 26-Aug-1999  
 C:Accession: S36451  
 R:Savelkoul, P.; de Kerf, D.; de Groot, L.; Willems, R.; Mooi, F.; van der Zeijst, B.; C  
 submitted to the EMBL Data Library, February 1993  
 A:Description: Characterization of fimX, a gene encoding a fimbrial subunit protein of B  
 A:Reference number: S36451  
 A:Accession: S36451  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-201 <SAV>  
 A:Cross-references: EMBL:X74118; NID:g396788; PIDN:CAA52215.1; PID:g396789  
 C:Genetics:  
 A:Gene: fimX  
 C:Superfamily: type 1 fimbrial protein

Query Match 1.5%; Score 8; DB 2; Length 201;  
 Best Local Similarity 100.0%; Pred. No. 5.2;  
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 4 LLGAALAG 11  
 |||||  
 Db 7 LLGAALAG 14

## RESULT 11

S01929  
 fimbrial protein precursor - Bordetella pertussis  
 C:Species: Bordetella pertussis  
 C:Date: 18-Oct-1989 #sequence\_revision 18-Oct-1989 #text\_change 26-Aug-1999  
 C:Accession: S01929  
 R:Pedroni, P.; Riboli, B.; de Ferra, F.; Grandi, G.; Toma, S.; Arico, B.; Rappuoli, R.  
 Mol. Microbiol. 2, 539-543, 1988  
 A:Title: Cloning of a novel pilin-like gene from Bordetella pertussis: homology to the f  
 A:Reference number: S01929; MUID:89013896; PMID:2902506  
 A:Accession: S01929  
 A:Molecule type: DNA  
 A:Residues: 1-201 <PDP>  
 A:Cross-references: EMBL:Y00556; NID:g39750; PIDN:CAA68634.1; PID:g39751  
 C:Genetics:  
 A:Gene: fimX  
 C:Superfamily: type 1 fimbrial protein  
 F,1-21/Domain: signal sequence #status predicted <SIG>  
 F,22-201/Product: fimbrial protein #status predicted <MAT>

Query Match 1.5%; Score 8; DB 2; Length 201;  
 Best Local Similarity 100.0%; Pred. No. 5.2;  
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 4 LLGAALAG 11  
 |||||  
 Db 7 LLGAALAG 14

## RESULT 12

F64390  
 coenzyme F420 hydrogenase (EC 1.12.99.1) gamma chain - Methanococcus jannaschii  
 N:Alternate names: coenzyme F420-reducing hydrogenase, gamma subunit  
 C:Species: Methanococcus jannaschii  
 C:Date: 13-Sep-1996 #sequence\_revision 13-Sep-1996 #text\_change 02-Aug-2002  
 C:Accession: F64390  
 R:Bult, C.J.; White, O.; Olsen, G.J.; Zhou, L.; Fleischmann, R.D.; Sutton, G.G.; Blake,  
 ; Reich, C.I.; Overbeek, R.; Kirkness, E.F.; Weinstock, K.G.; Merrick, J.M.; Glodek, A.;  
 raon, J.D.; Sadow, P.W.; Hanna, M.C.; Cotton, M.D.; Roberts, K.M.; Hurs, M.A.  
 Science 273, 1058-1073, 1996

A:Authors: Kaine, B.P.; Borodovsky, M.; Klenk, H.P.; Fraser, C.M.; Smith, H.O.; Woese, C  
 A:Title: Complete genome sequence of the methanogenic archaeon, Methanococcus jannaschii  
 A:Reference number: A64300; MUID:96337999; PMID:8688087  
 A:Accession: F64390

A:Status: preliminary; nucleic acid sequence not shown; translation not shown  
 A:Molecule type: DNA  
 A:Residues: 1-216 <BUL>  
 A:Cross-references: GB:U67519; GB:L77117; NID:gi591436; PIDN:AAB98722.1; PID:gi591442;  
 C:Genetics:

A:Map position: REV659511-658861  
 C:Superfamily: coenzyme F420 hydrogenase gamma chain; ferredoxin 2[4Fe-4S] homology  
 C:Keywords: oxidoreductase  
 F:162-211/Domain: ferredoxin 2[4Fe-4S] homology <FER3>

Query Match 1.5%; Score 8; DB 2; Length 216;  
 Best Local Similarity 100.0%; Pred. No. 5.6;  
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 101 KEIVDSYL 108  
 |||||  
 Db 117 KEIVDSYL 124

## RESULT 13

B90841  
 probable ABC transporter permease protein ECs1698 [imported] - Escherichia coli (strain  
 C:Species: Escherichia coli  
 C:Date: 18-Jul-2001 #sequence\_revision 18-Jul-2001 #text\_change 18-Jul-2001  
 C:Accession: B90841  
 R:Hayashi, T.; Makino, K.; Ohnishi, M.; Kurokawa, K.; Ishii, K.; Yokoyama, K.; Han, C.G.  
 Gasawara, N.; Yasunaga, T.; Kuhara, S.; Shiba, T.; Hattori, M.; Shinagawa, H.  
 DNA Res. 8, 11-22, 2001  
 A:Title: Complete genome sequence of enterohemorrhagic Escherichia coli O157:H7 and geno  
 A:Reference number: A99629; MUID:21156231; PMID:11258796

A:Accession: B90841  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-326 <HAY>  
 A:Cross-references: GB:BA000007; PIDN:BA835121.1; PID:gl3361163; GSPDB:GN00154  
 A:Experimental source: strain O157:H7, substrain RMD 0509952  
 C:Genetics:  
 A:Gene: ECs1698

Query Match 1.5%; Score 8; DB 2; Length 326;  
 Best Local Similarity 100.0%; Pred. No. 8.2;  
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 4 LLGAALAG 11  
 |||||  
 Db 63 LLGAALAG 70

## RESULT 14

A85699  
 probable ABC transporter permease protein Z1965 [imported] - Escherichia coli (strain O1  
 C:Species: Escherichia coli  
 C:Date: 16-Feb-2001 #sequence\_revision 16-Feb-2001 #text\_change 27-Nov-2001  
 C:Accession: A85699  
 R:Perna, N.T.; Plunkett III, G.; Burland, V.; Mau, B.; Glasner, J.D.; Rose, D.J.; Mayhew  
 iller, L.; Grotbeck, E.J.; Davis, N.W.; Lim, A.; Dimalanta, E.; Potamousis, K.; Apodaca,

Nature 409, 529-533, 2001  
A:Title: Genome sequence of enterohemorrhagic Escherichia coli O157:H7.  
A:Reference number: A85480; MUID:21074935; PMID:11206551  
A:Accession: A85699  
A>Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-326 <STO>  
A:Cross-references: GB:AE005174; NID:g12514905; PIDN:AGS6053.1; GSPDB:GN00145; UMGF:Z19  
A:Experimental source: strain O157:H7, substrain EDL933  
C:Genetics:  
A:Gene: Z1965

Query Match 1.5%; Score 8; DB 2; Length 326;  
Best Local Similarity 100.0%; Pred. No. 8.2;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 LLGAALAG 11  
|||||||  
Db 63 LLGAALAG 70

## RESULT 15

AF3614  
probable undecaprenyl-phosphate alpha-N-acetylglucosaminyltransferase (EC 2.4.1.1-) [imp  
C:Species: Brucella melitensis  
C:Date: 01-Feb-2002 #sequence\_revision 01-Feb-2002 #text\_change 03-Feb-2003  
C:Accession: AF3614  
R:DelVecchio, V.G.; Kapatral, V.; Redkar, R.J.; Patra, G.; Mujer, C.; Los, T.; Ivanova,  
; Mazur, M.; Goltzman, E.; Selkov, E.; Elzer, P.H.; Hagius, S.; O'Callaghan, D.; Letes  
Proc. Natl. Acad. Sci. U.S.A. 99, 443-448, 2002  
A:Title: The genome sequence of the facultative intracellular pathogen Brucella melitens  
A:Reference number: AD3252; PMID:11756688  
A:Accession: AF3614  
A>Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-354 <KUR>  
A:Cross-references: GB:AE008918; PIDN:AAL54081.1; PID:g17985038; GSPDB:GN00191  
A:Experimental source: strain 16M  
C:Genetics:  
A:Gene: BMEI10839  
A:Map position: 11  
C:Keywords: glycosyltransferase; hexosyltransferase

Query Match 1.5%; Score 8; DB 2; Length 354;  
Best Local Similarity 100.0%; Pred. No. 8.8;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 LLGAALAG 11  
|||||||  
Db 213 LLGAALAG 220

Search completed: May 5, 2004, 13:48:28  
Job time : 17.405 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 5, 2004, 13:35:24 ; Search time 10.4395 Seconds

(without alignments)  
2583.672 Million cell updates/sec

Title: US-09-743-684a-1\_COPY\_7\_524

Perfect score: 518

Sequence: 1 LASLGAALAGPVGLKECT.....NTETAQCNAYEHCKRHVMN 518

Scoring table: OLIGO

Searched: 141681 seqs, 52070155 residues

Word size: 0

Total number of hits satisfying chosen parameters: 141681

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Database: SwissProt\_42:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	518	100.0	524	1	SAP_HUMAN
2	518	100.0	524	1	SAP_HUMAN
3	518	100.0	524	1	SAP_HUMAN
4	518	100.0	524	1	SAP_HUMAN
5	518	100.0	524	1	SAP_HUMAN
6	518	100.0	524	1	SAP_HUMAN
7	518	100.0	524	1	SAP_HUMAN
8	518	100.0	524	1	SAP_HUMAN
9	518	100.0	524	1	SAP_HUMAN
10	518	100.0	524	1	SAP_HUMAN
11	518	100.0	524	1	SAP_HUMAN
12	518	100.0	524	1	SAP_HUMAN
13	518	100.0	524	1	SAP_HUMAN
14	518	100.0	524	1	SAP_HUMAN
15	518	100.0	524	1	SAP_HUMAN
16	518	100.0	524	1	SAP_HUMAN
17	518	100.0	524	1	SAP_HUMAN
18	518	100.0	524	1	SAP_HUMAN
19	518	100.0	524	1	SAP_HUMAN
20	518	100.0	524	1	SAP_HUMAN
21	518	100.0	524	1	SAP_HUMAN
22	518	100.0	524	1	SAP_HUMAN
23	518	100.0	524	1	SAP_HUMAN
24	518	100.0	524	1	SAP_HUMAN
25	518	100.0	524	1	SAP_HUMAN
26	518	100.0	524	1	SAP_HUMAN
27	518	100.0	524	1	SAP_HUMAN
28	518	100.0	524	1	SAP_HUMAN
29	518	100.0	524	1	SAP_HUMAN
30	518	100.0	524	1	SAP_HUMAN
31	518	100.0	524	1	SAP_HUMAN
32	518	100.0	524	1	SAP_HUMAN
33	518	100.0	524	1	SAP_HUMAN

## ALIGNMENTS

RESULT 1	ID	SAP_HUMAN	STANDARD	PRT	524 AA
AC	P07602	P07602	P15793	P78538	P78541; P78546; P78547; P78558;
AC	Q92739	Q92740	Q92741	Q92742	
DT	01-APR-1988	(Rel. 07, Created)			
DT	01-APR-1990	(Rel. 14, Last sequence update)			
DT	10-OCT-2003	(Rel. 42, Last annotation update)			
DE	Sapasin B (Sphingolipid activator protein 1) (SAP-1) (Cerebroside sulfate activator) (CSact) (Dispersin) (Sulfatide) (Glucosylceramidase activator) (Sphingolipid activator protein 2) (SAP-2); Sapasin D (Protein C) (Component C).				
DE	PSAP.				
GN	Homo sapiens (Human).				
OS	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;				
OC	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.				
OX	NCBI_TaxID=9606;				
RN	[1]				
RP	SEQUENCE FROM N.A.				
RC	TISSUE=Liver;				
RX	MEDLINE=90129043; PubMed=2515150;				
RA	Roman E.G., Grabowski G.A.;				
RT	"Molecular cloning of a human co-beta-glucosidase cDNA: evidence that four sphingolipid hydrolase activator proteins are encoded by single genes in humans and rats.";				
RT	Genomics 5:486-492(1989).				
RL	[2]				
RP	SEQUENCE FROM N.A.				
RX	MEDLINE=89255151; PubMed=2498298;				
RA	Nakano T., Sandhoff K., Struemper J., Christomanou H., Suzuki K.;				
RT	"Structure of full-length cDNA coding for sulfatide activator, a co-beta-glucosidase and two other homologous proteins: two alternate forms of the sulfatide activator.";				
RT	J. Biochem. 105:152-154(1989).				
RL	[3]				
RP	SEQUENCE FROM N.A. (ISOROM SAP-MU-0).				
RC	TISSUE=Brain, Eye, and Skin;				
RX	MEDLINE=22388257; PubMed=12477932;				
RA	Klausner R.D., Collins F.S., Wagner L., Shennan C.M., Schuler G.D.,				
RA	Altschul S.F., Zeeberg B., Buettow K.H., Schaefer C.F., Bhat N.K.,				
RA	Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,				
RA	Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,				
RA	Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Schneetz T.E.,				
RA	Brownstein M.J., Usdin T.B., Toshiyuki S., Carrinui P., Prange C.,				
RA	Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,				
RA	Bosch S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,				
RA	Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hilyk S.W.,				
RA	Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,				
RA	Faney J., Helton E., Kerteman M., Madan A., Rodríguez S., Sanchez A.,				
RA	Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,				
RA	Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,				
RA	Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,				
RA	Butterfield V.S.N., Krzywinski M.I., Skalska U., Smallos D.E.,				

34	7	1.4	298	1	XERD_PSEAE	Q9hxx6 pseudomonas
35	7	1.4	298	1	XERD_PSEBK	O88mv0 pseudomonas
36	7	1.4	309	1	Y209_MYCPN	P75485 mycoplasma
37	7	1.4	313	1	MRAM_PSEAE	Q9hvx5 pseudomonas
38	7	1.4	344	1	PUR5_NEIMA	Q9jua2 neisseria m
39	7	1.4	344	1	PUR5_NEIMB	Q9j280 neisseria m
40	7	1.4	388	1	RGSK_HUMAN	O76081 homo sapien
41	7	1.4	396	1	CAPE_STAVU	P39855 streptyllococ
42	7	1.4	396	1	TCR3_ECOLI	P02981 escherichia
43	7	1.4	400	1	TCR9_PASMO	P51564 pasteurella
44	7	1.4	401	1	NIFS_ENTAG	O52069 enterobacte
45	7	1.4	436	1	PDK1_HUMAN	Q15118 homo sapien

RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;  
 RT "Generation and initial analysis of more than 15,000 full-length  
 RT human and mouse cDNA sequences.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).  
 RP SEQUENCE OF 59-125 AND 304-513 FROM N.A.  
 RC TISSUE=Brain;  
 RA MEDLINE=91192146; PubMed=2013321;  
 RA Holtschmidt H., Sandhoff K., Fuest W., Kwon H.Y., Schnabel D.,  
 RA Suzuki K.;  
 RT "The organization of the gene for the human cerebroside sulfate  
 RT activator protein.";  
 RL FEBS Lett. 280:267-270(1991).  
 RN [5]  
 RP SEQUENCE OF 164-524 FROM N.A.  
 RX MEDLINE=8806647; PubMed=2825202;  
 RA Dewi N.N., Wenger D.A., O'Brien J.S.;  
 RT "Nucleotide sequence of cloned cDNA for human sphingolipid activator  
 RT protein 1 precursor.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 84:8652-8656(1987).  
 RN [6]  
 RP PARTIAL SEQUENCE OF 60-142.  
 RX MEDLINE=89240739; PubMed=2717620;  
 RA Morimoto S., Martin B.M., Yamamoto Y., Kretz K.A., O'Brien J.S.,  
 RA Kishimoto Y.;  
 RT "Saposin A: second cerebroside activator protein.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 86:3389-3393(1989).  
 RN [7]  
 RP SEQUENCE OF 195-263 FROM N.A.  
 RX MEDLINE=86130593; PubMed=2868718;  
 RA Dewi N.N., Wenger D.A., Fujibayashi S., Donoviel M., Esch F.,  
 RA Hill F., O'Brien J.S.;  
 RT "Molecular cloning of the sphingolipid activator protein-1 (SAP-1),  
 RT the sulfate sulfate activator.";  
 RL Biochem. Biophys. Res. Commun. 134:989-994(1986).  
 RN [8]  
 RP SEQUENCE OF 195-274.  
 RC TISSUE=Kidney;  
 RX MEDLINE=91006165; PubMed=2209618;  
 RA Furst W., Schuber J., Machleidt W., Meyer H.E., Sandhoff K.;  
 RT "The complete amino-acid sequences of human ganglioside GM2 activator  
 RT protein and cerebroside sulfate activator protein.";  
 RL Eur. J. Biochem. 192:709-714(1990).  
 RN [9]  
 RP SEQUENCE OF 195-274.  
 RX MEDLINE=89207118; PubMed=3242555;  
 RA Kleinschmidt T., Christomanou H., Braunitzer G.;  
 RT "Complete amino-acid sequence of the naturally occurring A2 activator  
 RT protein for enzymic sphingomyelin degradation: identity to the  
 RT sulfate activator protein (SAP-1).";  
 RL Biol. Chem. Hoppe-Seyler 369:1361-1365(1988).  
 RN [10]  
 RP SEQUENCE OF 311-390.  
 RX MEDLINE=88163077; PubMed=3442600;  
 RA Kleinschmidt T., Christomanou H., Braunitzer G.;  
 RT "Complete amino-acid sequence and carbohydrate content of the  
 RT naturally occurring glycosylceramide activator protein (A1 activator)  
 RT absent from a new human Gaucher disease variant.";  
 RL Biol. Chem. Hoppe-Seyler 368:1571-1578(1987).  
 RN [11]  
 RP SEQUENCE OF 407-484.  
 RX MEDLINE=89000190; PubMed=3048308;  
 RA Furst W., Machleidt W., Sandhoff K.;  
 RT "The precursor of sulfate activator protein is processed to three  
 RT different proteins.";  
 RL Biol. Chem. Hoppe-Seyler 369:317-328(1988).  
 RN [12]  
 RP PARTIAL SEQUENCE OF 405-484.  
 RX MEDLINE=89025876; PubMed=2845979;  
 RA Morimoto S., Martin B.M., Kishimoto Y., O'Brien J.S.;  
 RT "Saposin D: a sphingomyelinase activator.";  
 RL Biochem. Biophys. Res. Commun. 156:403-410(1988).  
 RN [13]  
 RP SEQUENCE OF 17-26.  
 RC TISSUE=Milk;  
 RX MEDLINE=92068206; PubMed=1958198;  
 RA Kondoh K., Hienno T., Sano A., Kakimoto Y.;  
 RT "Isolation and characterization of prosaposin from human milk.";  
 RL Biochem. Biophys. Res. Commun. 181:286-292(1991).  
 RN [14]  
 RP PARTIAL SEQUENCE (SAPOSIN B), AND STRUCTURE OF CARBOHYDRATES.  
 RC TISSUE=urine;  
 RX MEDLINE=20032116; PubMed=10562467;  
 RA Fluharty A.L., Lombardo C., Louis A., Stevens R.L., Whitelegge J.P.,  
 RA Waring A.V., To T., Fluharty C.B., Faull K.F.;  
 RT "Preparation of the cerebroside sulfate activator (CSact or saposin B)  
 RT from human urine.";  
 RL Mol. Genet. Metab. 68:391-403(1999).  
 RN [15]  
 RP STRUCTURE OF CARBOHYDRATE ON ASN-215.  
 RX MEDLINE=2110404; PubMed=11180632;  
 RA Faull K.F., Johnson J., Kim M.J., To T., Whitelegge J.P.,  
 RA Stevens R.L., Fluharty C.B., Fluharty A.L.;  
 RT "Structure of the asparagine-linked sugar chains of porcine kidney and  
 RT human urine cerebroside sulfate activator protein.";  
 RL J. Mass Spectrom. 35:1416-1424(2000).  
 RN [16]  
 RP SAPOSIN D DISULFIDE BONDS.  
 RX MEDLINE=99337688; PubMed=10406958;  
 RA Tatti M., Salvioli R., Ciaffoni F., Pucci P., Andolfo A.,  
 RA Amoresano A., Vaccaro A.M.;  
 RT "Structural and membrane-binding properties of saposin D.";  
 RL Eur. J. Biochem. 263:486-494(1999).  
 RN [17]  
 RP SAPOSIN B DISULFIDE BONDS.  
 RX MEDLINE=22398398; PubMed=12510003;  
 RA Ann V.E., Faull K.F., Whitelegge J.P., Higginson J., Fluharty A.L.,  
 RA Prive G.G.;  
 RT "Expression, purification, crystallization, and preliminary X-ray  
 RT analysis of recombinant human saposin B.";  
 RL Protein Expr. Purif. 27:186-193(2003).  
 RN [18]  
 RP MASS SPECTROMETRY.  
 RC TISSUE=Urine;  
 RX MEDLINE=99441404; PubMed=10510427;  
 RA Faull K.F., Whitelegge J.P., Higginson J., To T., Johnson J.,  
 RA Kruchinsky A.N., Stranding K.G., Waring A.V., Stevens R.L.,  
 RA Fluharty C.B., Fluharty A.L.;  
 RT "Cerebroside sulfate activator protein (Saposin B): chromatographic  
 RT and electrospray mass spectrometric properties.";  
 RL J. Mass Spectrom. 34:1040-1054(1999).  
 RN [19]  
 RP X-RAY CRYSTALLOGRAPHY (2.2 ANGSTROMS) OF 195-273, AND MUTAGENESIS OF  
 RP ILE-240.  
 RX MEDLINE=22406333; PubMed=12518053;  
 RA Ann V.E., Faull K.F., Whitelegge J.P., Fluharty A.L., Prive G.G.;  
 RT "Crystal structure of saposin B reveals a dimeric shell for lipid  
 RT binding.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 100:38-43(2003).  
 RN [20]  
 RP REVIEW ON MLD VARIANTS.  
 RX MEDLINE=95170731; PubMed=7866401;  
 RA Gieselmann V., Zlotogora J., Harris A., Wenger D.A., Morris C.P.;  
 RT "Molecular genetics of metachromatic leukodystrophy.";  
 RL Hum. Mutat. 4:233-242(1994).  
 RN [21]  
 RP VARIANT MLD ILE-217.  
 RX MEDLINE=90147748; PubMed=2302219;  
 RA Rafi M.A., Zhang X.-L., Degala G., Wenger D.A.;  
 RT "Detection of a point mutation in sphingolipid activator protein-1  
 RT leukodystrophy.";  
 RL Biochem. Biophys. Res. Commun. 166:1017-1023(1990).  
 RN [22]  
 RP SEQUENCE FROM N.A. AND VARIANT MLD ILE-217.  
 RX MEDLINE=90207231; PubMed=2320574;

Query Match 100.0%; Score 518; DB 1; Length 524;  
 Best Local Similarity 100.0%; Pred. No. 0;  
 Matches 518; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LASLIGALAGPVLKKECTRGSAVWCVKNTASDCGAKHCLQTVWKKPVPKSLPCDIC 60  
 DB 7 LASLIGALAGPVLKKECTRGSAVWCVKNTASDCGAKHCLQTVWKKPVPKSLPCDIC 66  
 QY 61 KDVTAAADMLKDNATEEELVLYLEKTCDELPKPKMSASCKEIVDSYLPVLIIDIKGMS 120  
 DB 67 KDVTAAADMLKDNATEEELVLYLEKTCDELPKPKMSASCKEIVDSYLPVLIIDIKGMS 126  
 QY 121 RRGVCSALNLCESIOKHAELNHOKLESNKIPELDMTEVAVAPPMANIPLLYPODPR 180  
 DB 127 RRGVCSALNLCESIOKHAELNHOKLESNKIPELDMTEVAVAPPMANIPLLYPODPR 186  
 QY 181 SKPQKNDGVQCQDCIQWVTDIQAVRTNSTFVQALVEHVEKCDRLGPMADICKNYIS 240  
 DB 187 SKPQKNDGVQCQDCIQWVTDIQAVRTNSTFVQALVEHVEKCDRLGPMADICKNYIS 246  
 QY 241 QYSEIAIOMMMHMOPEKICALVGFCDVEKEMQTLVPAKVASKVIVIPALIVEPIKHE 300  
 DB 247 QYSEIAIOMMMHMOPEKICALVGFCDVEKEMQTLVPAKVASKVIVIPALIVEPIKHE 306  
 QY 301 VPAKSDVYCEVCEPFLVKEVTYKLDINNKEKEILDAFDKMCSTLPSLSECOEVVDYGS 360  
 DB 307 VPAKSDVYCEVCEPFLVKEVTYKLDINNKEKEILDAFDKMCSTLPSLSECOEVVDYGS 366  
 QY 361 SILSILIEVSEPELVCSMLHLCSTGRLPALTVHTVQPKDGCCEVCKLVGLDNLNLEN 420  
 DB 367 SILSILIEVSEPELVCSMLHLCSTGRLPALTVHTVQPKDGCCEVCKLVGLDNLNLEN 426  
 QY 421 STKOEIILALEKGSFLPDYQKQCDQVAYEYEPVLIIEVWMDPSVCLIKIGACPSAH 480  
 DB 427 STKOEIILALEKGSFLPDYQKQCDQVAYEYEPVLIIEVWMDPSVCLIKIGACPSAH 486  
 QY 481 KPLGTEKICMGPSYWCNTETAAOCNAVEHKKRRHYVN 518  
 DB 487 KPLGTEKICMGPSYWCNTETAAOCNAVEHKKRRHYVN 524

RESULT 2  
 SAP\_BOVIN STANDARD; PRT; 525 AA.  
 ID SAP\_BOVIN STANDARD; PRT; 525 AA.  
 AC P26779; Q9NRG4; Rel. 23; Created)  
 DT 01-AUG-1992 (Rel. 40; Last sequence update)  
 DT 16-OCT-2001 (Rel. 40; Last sequence update)  
 DT 16-OCT-2001 (Rel. 40; Last sequence update)  
 DE Proactivator polypeptide precursor [Contains: Saposin A (Protein A);  
 Saposin B (Sphingolipid activator protein 1) (SAP-1) (Cerebroside  
 sulfate activator) (CSAct) (Dispersin) (Sulfatide/GM1 activator);  
 Saposin C (Co-beta-glucosidase) (A1 activator) (Glucosylceramidase  
 activator) (Sphingolipid activator protein 2) (SAP-2); Saposin D  
 (Protein C) (Component C)].  
 DE PSAP.  
 CN PSAP.  
 OS Bos taurus (Bovine).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidea;  
 OC Bovidae; Bovinae; Bos.  
 OC NCBI\_Taxid=9913;  
 RN [1]  
 RP SEQUENCE FROM N.A., AND VARIANTS.  
 RC TISSUE=Mammary gland;  
 RA Azuma N., Yoshida K.;  
 RT "RT-PCR cloning of bovine prosaposin.";  
 RL Submitted (JAN-2000) to the EMBL/GenBank/DBJ databases.  
 RN [2]  
 RP SEQUENCE OF 312-391.  
 RC TISSUE=Spleen;  
 RX MEDLINE=92207994; PubMed=1554743;  
 RA Sano A., Mizuno T., Kondoh K., Hino S.-I., Kakimoto Y.,  
 Morita N.;

RT "Saposin-C from bovine spleen; complete amino acid sequence and  
 RT relation between the structure and its biological activity.";  
 RL Biochim. Biophys. Acta 1120:75-80(1992).  
 CC -I- FUNCTION: The lysosomal degradation of sphingolipids takes place  
 CC by the sequential action of specific hydrolases. Some of these  
 CC enzymes require specific low-molecular mass, non-enzymic proteins:  
 CC the sphingolipids activator proteins (coproteins) (By similarity).  
 CC -I- FUNCTION: Saposin A and saposin C stimulate the hydrolysis of  
 CC glucosylceramide by beta-glucosylceramidase (EC 3.2.1.45) and  
 CC galactosylceramide by beta-galactosylceramidase (EC 3.2.1.46).  
 CC Saposin C apparently acts by combining with the enzyme and acidic  
 CC lipid to form an activated complex, rather than by solubilizing  
 CC the substrate.  
 CC -I- FUNCTION: Saposin B stimulates the hydrolysis of galacto-  
 CC cerebroside sulfate by arylsulfatase A (EC 3.1.6.8), GM1  
 CC gangliosides by beta-galactosidase (EC 3.2.1.23) and  
 CC globotriaosylceramide by alpha-galactosidase A (EC 3.2.1.22).  
 CC Saposin B forms a solubilizing complex with the substrates of the  
 CC sphingolipid hydrolases (By similarity).  
 CC -I- FUNCTION: Saposin D is a specific sphingomyelin phosphodiesterase  
 CC activator (EC 3.1.4.12) (By similarity).  
 CC -I- SUBUNIT: Saposin B is a homodimer (By similarity).  
 CC -I- SUBCELLULAR LOCATION: Lysosomal.  
 CC -I- PTM: This precursor is proteolytically processed to 4 small  
 CC peptides, which are similar to each other and are sphingolipid  
 CC hydrolase activator proteins (By similarity).  
 CC -I- SIMILARITY: Contains 2 saposin A-type domains.  
 CC -I- SIMILARITY: Contains 4 saposin B-type domains.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL: AB036791; BA95677.1; -  
 DR InterPro: IPR003119; SAPA.  
 DR InterPro: IPR007856; SAPB\_1.  
 DR InterPro: IPR008138; SAPB\_2.  
 DR InterPro: IPR008140; SAPB\_sub.  
 DR InterPro: IPR008373; Saposin.  
 DR InterPro: IPR008139; SaposinB.  
 DR Pfam: PF02199; SAPA; 2.  
 DR Pfam: PF05184; SAPB\_1; 4.  
 DR Pfam: PF03489; SAPB\_2; 4.  
 DR PRINTS: PR01797; SAPOSIN.  
 DR ProDom: PD001732; SAPB\_sub; 3.  
 DR SMART: SMO0162; SAPA; 2.  
 DR SMART: SMO0118; SAPB; 4.  
 KW Signal; Glycoprotein; Lysosome; Sphingolipid metabolism; Repeat.  
 FT SIGNAL 1 16  
 FT PROPEP 17 58  
 FT CHAIN 60 142  
 FT PROPEP 144 195  
 FT CHAIN 196 275  
 FT PROPEP 277 310  
 FT CHAIN 312 392  
 FT PROPEP 393 407  
 FT CHAIN 406 484  
 FT PROPEP 489 525  
 FT DOMAIN 21 54  
 FT DOMAIN 59 142  
 FT DOMAIN 194 276  
 FT DOMAIN 312 393  
 FT DOMAIN 406 487  
 FT DOMAIN 492 525  
 FT DISULFID 63 138  
 FT DISULFID 66 132  
 FT DISULFID 94 106  
 FT DISULFID 199 272  
 FT DISULFID 202 266

SAPOSIN A.  
 SAPOSIN B.  
 SAPOSIN C.  
 SAPOSIN D.  
 SAPOSIN-LIKE TYPE A 1.  
 SAPOSIN-LIKE TYPE B 1.  
 SAPOSIN-LIKE TYPE B 2.  
 SAPOSIN-LIKE TYPE B 3.  
 SAPOSIN-LIKE TYPE B 4.  
 SAPOSIN-LIKE TYPE A 2.  
 BY SIMILARITY.  
 BY SIMILARITY.  
 BY SIMILARITY.  
 BY SIMILARITY.  
 BY SIMILARITY.

```

FT DISULFID 231 242 BY SIMILARITY.
FT DISULFID 316 389 BY SIMILARITY.
FT DISULFID 319 383 BY SIMILARITY.
FT DISULFID 347 358 BY SIMILARITY.
FT DISULFID 410 483 BY SIMILARITY.
FT DISULFID 413 477 BY SIMILARITY.
FT DISULFID 441 452 BY SIMILARITY.
FT CARBOHYD 80 80 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 101 101 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 216 216 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 333 333 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 427 427 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT VARIANT 127 127 H -> R.
FT VARIANT 260 263 MODP -> IIRIR.
FT CONFLICT 317 317 E -> Q (IN REF. 2).
FT CONFLICT 367 367 R -> S (IN REF. 1).
SQ SEQUENCE 525 AA; 58120 MW; 293AF0FB9C4FA99 CRC64;

Query Match 5.8%; Score 30; DB 1; Length 525;
Best Local Similarity 100.0%; Pred. No. 8e-22;
Matches 30; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 122 PGEVCSALNCESLQGHAEINHQOLESN 151
DB 128 PGEVCSALNCESLQGHAEINHQOLESN 157

RESULT 3
SAP_PIG STANDARD; PRT; 80 AA.
AC P81A05;
DT 15-DEC-1998 (Rel. 37, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Saposin B (Cerebroside sulfate activator) (CS-ACT) (Non-specific
DE activator) (Sphingolipid activator protein 1) (SAP-1).
OS Sus scrofa (Pig).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Suidae; Suidae; Sus.
OC NCBI_TaxID=9823;
RN (1)
RP SEQUENCE OF 1-79.
RC TISSUE=Kidney;
RX MEDLINE=93229506; PubMed=8471613;
RA Stevens R.L., Pauli K.F., Conklin K.A., Green B.N., Fluharty A.L.;
RT "Porcine cerebroside sulfate activator: further structural
RT characterization and disulfide identification.";
RL Biochemistry 32:4051-4059(1993).
RN (2)
RP SEQUENCE OF 1-64.
RC TISSUE=Kidney;
RX MEDLINE=92222651; PubMed=1562358;
RA Fluharty A.L., Katona Z., Meek W.E., Frei K., Fowler A.V.;
RT "The cerebroside sulfate activator from pig kidney: purification and
RT molecular structure.";
RL Biochem. Med. Metab. Biol. 47:66-85(1992).
RN (3)
RP STRUCTURE OF CARBOHYDRATE ON ASN-21.
RX MEDLINE=21110404; PubMed=11180632;
RA Pauli K.F., Johnson J., Kim M.J., To T., Whitelegge J.P.,
RA Stevens R.L., Fluharty C.B., Fluharty A.L.;
RT "Structure of the asparagine-linked sugar chains of porcine kidney and
RT human urine cerebroside sulfate activator protein.";
RL J. Mass Spectrom. 35:1416-1424(2000).
RN (4)
RP MASS SPECTROMETRY.
RC TISSUE=Kidney;
RX MEDLINE=99441404; PubMed=10510427;
RA Pauli K.F., Whitelegge J.P., Higginson J., To T., Johnson J.,
RA Krutichinsky A.N., Standing K.G., Waring A.J., Stevens R.L.,
RA Fluharty C.B., Fluharty A.L.;
RT "Cerebroside sulfate activator protein (Saposin B): chromatographic
RT and electrospray mass spectrometric properties.";

```

```

RL J. Mass Spectrom. 34:1040-1054(1999).
CC -1- FUNCTION: Saposin B stimulates the hydrolysis of galacto-
CC cerebroside sulfate by arylsulfatase A (EC 3.1.6.8), GM1
CC gangliosides by beta-galactosidase (EC 3.2.1.23) and
CC globotriaosylceramide by alpha-galactosidase A (EC 3.2.1.22).
CC Saposin B forms a solubilizing complex with the substrates of the
CC sphingolipid hydrolases.
CC -1- SUBUNIT: Saposin B is a homodimer (By similarity).
CC -1- PTM: The one residue extended Saposin B-Val is only found in a
CC minority of the chains.
CC -1- SIMILARITY: Contains 1 saposin B-type domain.
DR GLYCOSULEDB; P81A05; -.
DR InterPro; IPR007856; SAPB_1.
DR InterPro; IPR008138; SAPB_2.
DR InterPro; IPR008373; Saposin.
DR InterPro; IPR008139; SaposinB.
DR Pfam; PF05184; SAPB_1; 1.
DR Pfam; PF03489; SAPB_2; 1.
DR PRINTS; PR01797; SAPOSIN.
DR SMART; SM00118; SAPB; 1.
DR SMART; SM00118; SAPB; 1.
DR Glycoprotein; Sphingolipid metabolism.
FT CHAIN 1 79 SAPOSIN B.
FT CHAIN 1 80 SAPOSIN B-VAL.
FT DOMAIN 1 80 SAPOSIN-LIKE TYPE B.
FT DISULFID 4 77
FT DISULFID 7 71
FT DISULFID 36 47
FT CARBOHYD 21 21
SQ SEQUENCE 80 AA; 8949 MW; E7BA249B3E789C CRC64;
/FTID=CAR_000177.

Query Match 4.2%; Score 22; DB 1; Length 80;
Best Local Similarity 100.0%; Pred. No. 1.6e-14;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 235 CKNYISQYSEIATQMMHMQPK 256
DB 47 CKNYISQYSEIATQMMHMQPK 68

RESULT 4
SAP_RAT STANDARD; PRT; 554 AA.
ID SAP_RAT
AC P10560; O62841; O64190;
DT 01-JUL-1989 (Rel. 11, Created)
DT 01-JUL-1989 (Rel. 11, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE Sulfated glycoprotein 1 precursor (SGP-1) (Prosaposin).
DE PSAP OR SGP1.
GN PSAP OR SGP1.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OC NCBI_TaxID=10116;
RN (1)
RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
RC TISSUE=Sertoli cells;
RX MEDLINE=8900647; PubMed=3048385;
RA Collard M.W., Sylvester S.R., Tsunura J.K., Griswold M.D.;
RT "Biosynthesis and molecular cloning of sulfated glycoprotein 1
RT secreted by rat Sertoli cells: sequence similarity with the
RT 70-kilodalton precursor to sulfatide/GM1 activator.";
RL Biochemistry 27:4557-4564(1988).
RN (2)
RP SEQUENCE FROM N.A.
RC TISSUE=Testis;
RX MEDLINE=96128541; PubMed=8573994;
RA Morales C.R., El-Alfy M., Zhao Q., Igodou S.A.;
RT "Molecular role of sulfated glycoprotein-1 (SGP-1/prosaposin) in
RT Sertoli cells.";
RL Histol. Histopathol. 10:1023-1034(1995).
RN (3)
RP SEQUENCE FROM N.A.
RC TISSUE=Testis;

```

```

RX MEDLINE=96175245; PubMed=8601692;
RA Morales C.R., El-Alfy M., Zhao Q., Igddoura S.A.;
RT "Expression and tissue distribution of rat sulfated glycoprotein-1
RT (prosaposin).";
RL J. Histochem. Cytochem. 44:327-337(1996).
CC -1- SUBCELLULAR LOCATION: Extracellular.
CC -1- SIMILARITY: Contains 2 saposin A-type domains.
CC -1- SIMILARITY: Contains 4 saposin B-type domains.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; M19936; AAA42136.1; -
CC EMBL; S81353; AAB36042.2; -
CC EMBL; S81373; AAB36233.2; -
CC PIR; A28716; A28716.
CC InterPro; IPR003119; SAPA.
CC InterPro; IPR007856; SAPB_1.
CC InterPro; IPR008138; SAPB_2.
CC InterPro; IPR008140; SAPB_sub.
CC InterPro; IPR008373; Saposin.
CC InterPro; IPR008139; SaposinB.
CC Pfam; PF02199; SAPA; 2.
CC Pfam; PF05184; SAPB_1; 3.
CC Pfam; PF03489; SAPB_2; 4.
CC PRINTS; PR01797; SAPOSIN.
CC PRODOM; PD001732; SAPB_sub; 3.
CC SMART; SM00162; SAPA; 2.
CC SMART; SM00118; SAPB; 4.
CC Signal; Glycoprotein; Repeat.
CC -----
FT CHAIN 1 554 SULFATED GLYCOPROTEIN 1.
FT SIGNAL 1 16
FT DOMAIN 17 54 SAPOSIN-LIKE TYPE A 1.
FT DOMAIN 21 54 SAPOSIN-LIKE TYPE B 1.
FT DOMAIN 61 138 SAPOSIN-LIKE TYPE B 2.
FT DOMAIN 193 274 SAPOSIN-LIKE TYPE B 3.
FT DOMAIN 310 391 SAPOSIN-LIKE TYPE B 4.
FT DOMAIN 435 516 SAPOSIN-LIKE TYPE B 5.
FT DOMAIN 521 554 SAPOSIN-LIKE TYPE A 2.
FT DOMAIN 63 138 BY SIMILARITY.
FT DISULFID 66 132 BY SIMILARITY.
FT DISULFID 94 106 BY SIMILARITY.
FT DISULFID 197 270 BY SIMILARITY.
FT DISULFID 200 264 BY SIMILARITY.
FT DISULFID 229 240 BY SIMILARITY.
FT DISULFID 314 387 BY SIMILARITY.
FT DISULFID 317 381 BY SIMILARITY.
FT DISULFID 345 356 BY SIMILARITY.
FT DISULFID 439 512 BY SIMILARITY.
FT DISULFID 442 506 BY SIMILARITY.
FT DISULFID 470 481 BY SIMILARITY.
FT CARBOHYD 80 80 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 214 214 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 331 351 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 456 456 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CONFLICT 115 115 P -> L (IN REF. 2 AND 3).
FT CONFLICT 299 299 D -> V (IN REF. 2).
FT CONFLICT 462 527 I -> R (IN REF. 3).
FT CONFLICT 527 527 W -> R (IN REF. 3).
FT CONFLICT 536 536 S -> M (IN REF. 3).
SQ SEQUENCE 554 AA; 61123. MW; DFE3F3A3A0520C6B CRC64;

```

```

RESULT 5
SAP_MOUSE STANDARD; PRT; 557 AA.
ID SAP_MOUSE 061207; 060861; 064006; 064219;
AC 15-DEC-1998 (Rel. 37, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE Sulfated glycoprotein 1 precursor (SGP-1) (Prosaposin).
GN PSAP OR SGP1.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxId=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Liver;
RX MEDLINE=92272718; PubMed=1590788;
RA Tsuda M., Sakiyama T., Endo H., Kitagawa T.;
RT "The primary structure of mouse saposin.";
RL Biochem. Biophys. Res. Commun. 184:1266-1272(1992).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=94272317; PubMed=8003952;
RA Sprecher-Levy H., Orr-Urtreger A., Lona P., Horowitz M.;
RT "Murine prosaposin: expression in the reproductive system of a gene
RT implicated in human genetic disease.";
RL Cell. Mol. Biol. 40:233-233(1994).
RN [3]
RP SEQUENCE FROM N.A.
RX MEDLINE=96084310; PubMed=8565332;
RA Cao Q.P., Crain M.R.;
RT "Expression of SGP-1 mRNA in preimplantation mouse embryos.";
RL Dev. Genet. 17:263-271(1995).
RN [4]
RP SEQUENCE FROM N.A.
RC STRAIN=BALE/c; TISSUE=Liver;
RA Zhao Q.Q., Hay N.N., Morales C.R.;
RL Submitted (MAY-1996) to the EMBL/GenBank/DBJ databases.
CC -1- SUBCELLULAR LOCATION: Extracellular.
CC -1- SIMILARITY: Contains 2 saposin A-type domains.
CC -1- SIMILARITY: Contains 4 saposin B-type domains.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; S36200; AAB22175.1; -
CC EMBL; S71616; AAB31059.1; -
CC EMBL; U27340; AAB92567.1; -
CC EMBL; U57999; AAB02595.1; -
CC PIR; JH0604; JH0604.
CC MGD; MGI:97783; Psap.
CC InterPro; IPR003119; SAPA.
CC InterPro; IPR007856; SAPB_1.
CC InterPro; IPR008138; SAPB_2.
CC InterPro; IPR008140; SAPB_sub.
CC InterPro; IPR008373; Saposin.
CC Pfam; PF02199; SAPA; 2.
CC Pfam; PF05184; SAPB_1; 4.
CC Pfam; PF03489; SAPB_2; 4.
CC PRINTS; PR01797; SAPOSIN.
CC PRODOM; PD001732; SAPB_sub; 3.
CC SMART; SM00162; SAPA; 2.
CC SMART; SM00118; SAPB; 4.
CC Signal; Glycoprotein; Repeat.
FT SIGNAL 1 16 BY SIMILARITY.

```

```

Oy 425 ETLAALKGCSPFPYQKQCD 446
Db 461 ETLAALKGCSPFPYQKQCD 482

```

```

Query Match 4.2%; Score 22; DB 1; Length 554;
Best Local Similarity 100.0%; Pred. No. 8.9e-14;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```
FT CHAIN 17 557 SULFATED GLYCOPROTEIN 1.
FT DOMAIN 21 54 SAPOSIN-LIKE TYPE A 1.
FT DOMAIN 59 142 SAPOSIN-LIKE TYPE B 1.
FT DOMAIN 193 277 SAPOSIN-LIKE TYPE B 2.
FT DOMAIN 313 394 SAPOSIN-LIKE TYPE B 3.
FT DOMAIN 438 519 SAPOSIN-LIKE TYPE B 4.
FT DOMAIN 524 557 SAPOSIN-LIKE TYPE A 2.
FT DISULFID 63 138 BY SIMILARITY.
FT DISULFID 66 132 BY SIMILARITY.
FT DISULFID 94 106 BY SIMILARITY.
FT DISULFID 197 273 BY SIMILARITY.
FT DISULFID 200 267 BY SIMILARITY.
FT DISULFID 229 240 BY SIMILARITY.
FT DISULFID 317 390 BY SIMILARITY.
FT DISULFID 320 384 BY SIMILARITY.
FT DISULFID 348 359 BY SIMILARITY.
FT DISULFID 442 515 BY SIMILARITY.
FT DISULFID 445 509 BY SIMILARITY.
FT CARBOHYD 473 484 BY SIMILARITY.
FT CARBOHYD 80 80 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 214 214 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 334 334 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 459 459 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CONFLICT 83 83 O -> E (IN REF. 2).
FT CONFLICT 158 158 I -> V (IN REF. 3).
FT CONFLICT 160 160 MISSING (IN REF. 3).
FT CONFLICT 171 172 MS -> SA (IN REF. 2).
FT CONFLICT 244 244 V -> L (IN REF. 2).
FT CONFLICT 254 254 M -> I (IN REF. 3).
FT CONFLICT 255 255 L -> W (IN REF. 2).
FT CONFLICT 260 262 MISSING (IN REF. 3).
FT CONFLICT 307 307 N -> D (IN REF. 2).
FT CONFLICT 322 322 F -> L (IN REF. 2).
FT CONFLICT 345 350 AL -> GV (IN REF. 1).
FT CONFLICT 367 367 G -> D (IN REF. 3).
FT CONFLICT 370 370 L -> Q (IN REF. 2).
FT CONFLICT 373 373 I -> S (IN REF. 3).
FT CONFLICT 391 391 A -> T (IN REF. 3).
FT CONFLICT 393 393 R -> L (IN REF. 3).
FT CONFLICT 406 406 A -> R (IN REF. 2 AND 3).
FT CONFLICT 430 430 P -> R (IN REF. 2).
FT CONFLICT 445 445 C -> F (IN REF. 4).
FT CONFLICT 448 448 L -> P (IN REF. 4).
SQ SEQUENCE 557 AA; 61422 MW; 134593E2049935E CMC64;

Query Match 4.2%; Score 22; DB 1; Length 557;
Best Local Similarity 100.0%; Pred. No. 9e-14;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 425 EILALEKGSFLPPYOKCD 446
Db 464 EILALEKGSFLPPYOKCD 485

RESULT 6
SAP_CHICK STANDARD; PRT; 518 AA.
ID SAP_CHICK
AC 013035;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Proactivator polypeptide precursor [Contains: Saposin A; Saposin B;
DE Saposin C; Saposin D].
GN PSAP.
OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OX NCBI_TaxID=9031;
RN [1]
RP SEQUENCE FROM N.A., AND SEQUENCE OF 194-203.
RC TISSUE=Brain, and Liver;
RX MEDLINE=98129745; PubMed=9461526;
```

```
RA Azuma N., Seo H.-C., Lie O., Fu O., Gould R.M., Hiraiwa M., Burt D.W.,
RA Paton I.R., Morrice D.R., O'Brien J.S., Kishimoto Y.,
RT "Cloning, expression and map assignment of chicken prosaposin.";
RL Biochem. J. 330:321-327(1998).
RN [2]
RP SEQUENCE FROM N.A.
RA Altman N., Horowitz M.;
RL Submitted (NOV-1998) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: The lysosomal degradation of sphingolipids takes place
CC by the sequential action of specific hydrolases. Some of these
CC enzymes require specific low-molecular mass, non-enzymic proteins:
CC the sphingolipids activator proteins (coproteins) (By similarity).
CC -1- FUNCTION: Saposin A and saposin C stimulate the hydrolysis of
CC glucosylceramide by beta-glucosylceramidase (EC 3.2.1.45) and
CC galactosylceramide by beta-galactosylceramidase (EC 3.2.1.46).
CC Saposin C apparently acts by combining with the enzyme and acidic
CC lipid to form an activated complex, rather than by solubilizing
CC the substrate (By similarity).
CC -1- FUNCTION: Saposin B stimulates the hydrolysis of galacto-
CC cerebroside sulfate by arylsulfatase A (EC 3.1.6.8), GMI
CC globotriaosylceramide by alpha-galactosidase A (EC 3.2.1.22).
CC Saposin B forms a solubilizing complex with the substrates of the
CC sphingolipid hydrolases (By similarity).
CC -1- FUNCTION: Saposin D is a specific sphingomyelin phosphodiesterase
CC activator (EC 3.1.4.12) (By similarity).
CC -1- SUBUNIT: Saposin B is a homodimer (By similarity).
CC -1- PTM: This precursor is proteolytically processed to 4 small
CC peptides, which are similar to each other and are sphingolipid
CC hydrolase activator proteins (By similarity).
CC -1- SIMILARITY: Contains 2 saposin A-cyste domains.
CC -1- SIMILARITY: Contains 4 saposin B-cyste domains.
CC -----
CC This SWISS-PROT entry is copyrighted. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation-
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sb.ch/announce/
CC or send an email to license@isb-sb.ch).
CC -----
DR EMBL; AB003471; BAA19914.1; -
DR EMBL; AF108656; BAA05899.1; -
DR InterPro; IPR003119; SAPA.
DR InterPro; IPR007856; SAPB_1.
DR InterPro; IPR008138; SAPB_2.
DR InterPro; IPR008140; SAPB_sub.
DR InterPro; IPR008173; Saposin.
DR Pfam; PF02199; SAPA; 2.
DR Pfam; PF03489; SAPB_1; 4.
DR Pfam; PF03489; SAPB_2; 4.
DR PRINTS; PR01757; SAPOSIN.
DR ProDom; PD001732; SapB_sub; 3.
KW Signal; Glycoprotein; Lysosome; sphingolipid metabolism; Repeat;
KW GM2-gangliosidosis.
FT SIGNAL 1 17 POTENTIAL.
FT PROPEP 18 60
FT CHAIN 61 143 SAPOSIN A.
FT PROPEP 145 193
FT CHAIN 194 276 SAPOSIN B.
FT PROPEP 278 305
FT CHAIN 307 387 SAPOSIN C.
FT PROPEP 389 398
FT CHAIN 399 480
FT PROPEP 482 518 SAPOSIN D.
FT DOMAIN 22 55
FT DOMAIN 60 143 SAPOSIN-LIKE TYPE A 1.
FT DOMAIN 193 277 SAPOSIN-LIKE TYPE B 1.
FT DOMAIN 307 388 SAPOSIN-LIKE TYPE B 2.
FT DOMAIN 399 480 SAPOSIN-LIKE TYPE B 3.
FT DOMAIN 485 518 SAPOSIN-LIKE TYPE B 4.
FT DOMAIN 485 518 SAPOSIN-LIKE TYPE A 2.
```

```

FT DISULFID 64 139 BY SIMILARITY.
FT DISULFID 67 103 BY SIMILARITY.
FT DISULFID 95 107 BY SIMILARITY.
FT DISULFID 197 273 BY SIMILARITY.
FT DISULFID 200 267 BY SIMILARITY.
FT DISULFID 229 240 BY SIMILARITY.
FT DISULFID 311 384 BY SIMILARITY.
FT DISULFID 314 378 BY SIMILARITY.
FT DISULFID 342 353 BY SIMILARITY.
FT DISULFID 403 476 BY SIMILARITY.
FT DISULFID 406 470 BY SIMILARITY.
FT DISULFID 434 445 BY SIMILARITY.
FT CARBOHYD 81 81 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 214 214 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 328 328 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 420 420 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CONFLICT 94 94 R -> T (IN REF. 2).
FT CONFLICT 486 486 E -> D (IN REF. 2).
SQ SEQUENCE 518 AA; 57601 MW; B803000E891C3963 CRC64;

```

Query Match  
 Best Local Similarity 100.0%; Score 12; DB 1; Length 518;  
 Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 100 CKEIVDSYLPVI 111  
 |||||  
 Db 107 CKEIVDSYLPVI 118

## RESULT 7

SAP\_CAVPO STANDARD; PRT; 81 AA.

ID SAP\_CAVPO

AC P20097;

DT 01-FEB-1991 (Rel. 17, Last sequence update)

DT 01-FEB-1991 (Rel. 17, Last sequence update)

DT 16-OCT-2001 (Rel. 40, Last annotation update)

DE Saposin C (CO-beta-glucosidase) (Glucosylceramidase activator)

OS (Sphingolipid activator protein 2) (SAP-2).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Euteleostomi; Rodentia; Hystriognath; Cavidae; Cavia.

NCBI\_TaxID=10141;

RN [1]

RP SEQUENCE.

RC TISSUE=Liver;

RA MEDLINE=89066787; PubMed=3198642;

RA Sano A., Radin N.S., Johnson L.L., Tarr G.E.;

RT "The activator protein for glucosylceramidase beta-glucosidase from

guinea pig liver. Improved isolation method and complete amino acid

sequence."

RT J. Biol. Chem. 263:19597-19601(1988).

CC -I- FUNCTION: Saposin A and saposin C stimulate the hydrolysis of

glucosylceramide by beta-glucosylceramidase (EC 3.2.1.45) and

galactosylceramide by beta-galactosylceramidase (EC 3.2.1.46).

CC Saposin C apparently acts by combining with the enzyme and acidic

lipid to form an activated complex, rather than by solubilizing

the substrate.

CC -I- SIMILARITY: Contains 1 saposin B-type domain.

DR PIR: A32026; A32026.

DR HSSP: P42210; 10DM.

DR InterPro: IPR007856; SAPB\_1.

DR InterPro: IPR008138; SAPB\_2.

DR InterPro: IPR008140; SAPB\_sub.

DR InterPro: IPR008373; Saposin.

DR InterPro: IPR008139; SaposinB.

DR Pfam: PF05184; SAPB\_1; 1.

DR Pfam: PF03489; SAPB\_2; 1.

DR PRINTS: PR01797; SAPOSIN.

DR ProDom: PD001732; SAPB\_sub; 1.

DR SMART: SM00116; SAPB\_1.

DR GlycoProtein: Sphingolipid metabolism.

KM SAPOSIN-LIKE TYPE B.

FT DOMAIN 1 81

FT CARBOHYD 22 22 N-LINKED (GLCNAC. . .).

```

FT DISULFID 5 78 BY SIMILARITY.
FT DISULFID 8 72 BY SIMILARITY.
FT DISULFID 36 47 BY SIMILARITY.
SQ SEQUENCE 81 AA; 8852 MW; E564CE1F0A292596 CRC64;

```

Query Match  
 Best Local Similarity 100.0%; Score 9; DB 1; Length 81;  
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 351 COEVDITVG 359  
 |||||  
 Db 47 COEVDITVG 55

## RESULT 8

ID HS23\_DROME STANDARD; PRT; 186 AA.

AC P02516; OSVSX5;

DT 21-JUL-1986 (Rel. 01, Last sequence update)

DT 01-NOV-1990 (Rel. 16, Last sequence update)

DT 10-OCT-2003 (Rel. 42, Last annotation update)

DE Heat shock protein 23.

GN HSP23 OR CG4463.

OS Drosophila melanogaster (Fruit fly).

OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;

OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;

OC Ephydroidea; Drosophilidae; Drosophila.

NCBI\_TaxID=7227;

RN [1]

RP SEQUENCE FROM N.A.

RA MEDLINE=83189140; PubMed=6302284;

RA Southgate R., Ayme A., Voellmy R.;

RT "Nucleotide sequence analysis of the Drosophila small heat shock gene

cluster at locus 67B."

RT J. Mol. Biol. 165:35-57(1983).

RN [2]

RP SEQUENCE FROM N.A.

RA MEDLINE=82248004; PubMed=6285380;

RA Ingolia T.D., Craig E.A.;

RT "Four small Drosophila heat shock proteins are related to each other

and to mammalian alpha-crystallin."

RT Proc. Natl. Acad. Sci. U.S.A. 79:2360-2364(1982).

RN [3]

RP SEQUENCE FROM N.A.

RC STRAIN=Berkley;

RA MEDLINE=20196006; PubMed=10731132;

RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,

RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,

RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,

RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,

RA Brandon R.C., Rogers Y.-H.C., Blazey R.G., Chame M., Pfeiffer B.D.,

RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,

RA Abilaj J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,

RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu I., Beasley E.M.,

RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,

RA Borkovec D., Botchan M.R., Bouck J., Brockstein P., Brothier P.,

RA Burris K.C., Busam D.A., Butler H., Cadieu B., Center A., Chandra I.,

RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies S.M.,

RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,

RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,

RA Durdin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,

RA Foster C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glaser K.,

RA Glisler A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,

RA Harris N.L., Harvey D.A., Helman T.J., Hernandez J.R., Houck J.,

RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,

RA Jalaal M., Kalush F., Karpen G.H., Ke Z., Kemson J.A., Ketchum K.A.,

RA Kimmel B.E., Kodira C., Kraft C., Kravitz S., Kuip D., Lai Z.,

RA Laekko P., Lei Y., Levitsky A.A., Li J.H., Li Z., Liang Y., Lin X.,

RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,

RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,

RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,

RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Paclet J.M.,

RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,



```

RA Reiner K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA Spler E., Spreading A.C., Stapleton M., Strong R., Sun E.,
RA Svetske R., Tecor C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Wassarman D.A., Weinstein G.M., Weisenbach J.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
"the genome sequence of Drosophila melanogaster";
Science 287:2185-2195(2000).
[4]
SEQUENCE FROM N.A.
RA STRAIN=Berkeley; TISSUE=Embryo;
RX MEDLINE=22426066; PubMed=12531569;
RA Stapleton M., Carlson J.W., Brokstein P., Yu C., Champe M.,
RA George R.A., Guarin H., Krommiller B., Pacleb J.M., Park S., Wan K.H.,
RA Rubin G.M., Celinker S.E.;
" A Drosophila full-length cDNA resource.";
RL Genome Biol. 3:RESEARCH0080.1-RESEARCH0080.8(2002).
CC -1- SMILARITY: Belongs to the small heat shock protein (HSP20)
CC family.
CC
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
EMBL: J01100; AAA28637.1; -
EMBL: X00210; CAA23494.1; -
EMBL: X03889; CAA27525.1; -
EMBL: AE003552; AAF50286.1; -
EMBL: AY061081; AAL28629.1; -
DR PIR; B20647; B20647.
DR FLYBase; FBgn0001224; Hsp23.
DR InterPro; IPR001436; Crystallin_alpha.
DR InterPro; IPR002068; Hsp20.
DR InterPro; IPR008978; Hsp20_chap.
DR Pfam; PF00011; HSP20_1.
DR PRINTS; PRO0299; ACRYSTALLIN.
DR PROSITE; PS01031; HSP20_1.
KW Heat shock; Multigene family.
FT CONFLICT 31 33 GNN -> RRI (IN REF. 2).
FT CONFLICT 81 81 K -> E (IN REF. 2).
FT CONFLICT 88 88 K -> G (IN REF. 2).
FT CONFLICT 95 95 L -> V (IN REF. 2).
FT CONFLICT 173 173 N -> S (IN REF. 2).
FT CONFLICT 182 182 N -> G (IN REF. 2).
SQ SEQUENCE 186 AA; 20629 MW; 3CEDDDA171955A0 CRC64;
Query Match
Best Local Similarity 1.5%; Score 8; DB 1; Length 186;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 166 MANIPULL 173
Db 1 MANIPULL 8
RESULT 9
ID EMFX BORPE STANDARD; PRT; 201 AA.
AC P09808;
DT 01-MAR-1989 (Rel. 10, Created)
DT 10-OCT-2003 (Rel. 42, Last sequence update)
DE 15-MAR-2004 (Rel. 43, Last annotation update)
DE Fimbrial protein fmx precursor (Pilin).
OS FTMX OR BP2674.
ON Bordetella pertussis.
CC Bacteria; Proteobacteria; Betaproteobacteria; Burkholderiales;

```

```

OC Alcaligenaceae; Bordetella.
OX NCBI_TaxID=520;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=SA1;
RX MEDLINE=89013896; PubMed=2902506;
RA Pedroni P., Riboldi B., de Ferra F., Grandi G., Toma S., Arico B.,
RT Rappunioi R.;
RT "Cloning of a novel pilin-like gene from Bordetella pertussis:
RT homology to the fim2 gene.";
RT Mol. Microbiol. 2:539-543(1988).
RL [2]
RM
RP SEQUENCE FROM N.A.
RC STRAIN=Tohama I. / ATCC BAA-589 / NCTC 13251;
RX MEDLINE=22827954; PubMed=12910271;
RA Parkhill J., Sepalitha M., Preston A., Murphy L.D., Thomson N.,
RA Harris D.E., Holden M.T.G., Churcher C.M., Bentley S.D., Mungall K.L.,
RA Achtman M., Atkin R., Baker S., Basham D., Baeson N., Cherevach I.,
RA Chillingworth T., Collins M., Cronin A., Davis P., Doggett J.,
RA Fellwell T., Goble A., Hamlin N., Hauser H., Holtroyd S., Jørgels K.,
RA Leather S., Moule S., Norberczak H., O'Neill S., Ormond D., Price C.,
RA Rabinovitch E., Rutter S., Sanders W., Saunders D., Seeger K.,
RA Sharp S., Shummonds M., Skelton J., Squares R., Squares S., Stevens K.,
RA Umlin L., Whitehead S., Barrill B.G., Maskell D.J.;
RT "Comparative analysis of the genome sequences of Bordetella pertussis,
RT Bordetella parapertussis and Bordetella bronchiseptica.";
RT Nat. Genet. 35:32-40(2003).
RL
RC -1- SUBCELLULAR LOCATION: Pili structure on the cell surface.
CC -1- DISEASE: Bordetella pertussis is the causative agent of whooping
CC cough. An essential step in the disease process is the attachment
CC of the bacteria to the ciliated epithelium of the respiratory
CC tract, enabling the organism to resist normal host-clearance
CC mechanisms. It is unclear which bacterial cell surface component
CC are responsible for adherence but the fimbriae of B.pertussis are
CC prime candidates for being involved in this process.
CC -1- SIMILARITY: BELONGS TO THE FIMB/PAPA FAMILY OF FIMBRIA PROTEINS.
CC
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC at the European Bioinformatics Institute. There are no restrictions on
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sdb.ch/announce/
CC or send an email to license@sdb.ch).
CC
CC EMBL: Y00556; CAA68634.1; -.
CC EMBL: A25464; CAA01770.1; -.
CC EMBL: BX640419; CAB42950.1; -.
CC PIR: S01929; S01929.
CC DR InterPro: IPR008966; Adhes_bact.
CC DR InterPro: IPR000259; Fimbrtrial.
CC DR Pfam: PF00419; Fimbrtrial.1.
CC DR Fimbrtrial; Signal; Complete proteome.
CC FT SIGNAL 1 21
CC FT CHAIN 22 201 FIMBRIAL PROTEIN FIMX.
CC FT DISULFID 37 79 PROBABLE. (IN REF. 1).
CC FT CONFLICT 165 166 AK -> GQ (IN REF. 1).
CC SQ SEQUENCE 201 AA; 21462 MW; F47670CC055B03f6 CRC64;
CC
CC Query Match 1.5%; Score 8; DB 1; Length 201;
CC Best Local Similarity 100.0%; Pred. No. 4.1;
CC Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
CC
CC QY 4 LIGATING 11
CC DB 7 LIGATING 14
CC
CC RESULT 10
CC ID Y726_METJA STANDARD; PRT; 216 AA.
CC AC Q58136;

```



DT 01-NOV-1997 (Rel. 35, Created)  
 DT 01-NOV-1997 (Rel. 35, Last sequence update)  
 DT 10-OCT-2003 (Rel. 42, Last annotation update)  
 DE Hypothetical protein MU0726.  
 GN MU0726.  
 OS Methanococcus jannaschii.  
 OC Archaeae; Euryarchaeota; Methanococci; Methanococcales;  
 CC Methanocaldococcaceae; Methanocaldococcus.  
 OX NCBI\_TaxID=2190;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=JAL-1 / DSM 2661 / ATCC 43067;  
 RX MEDLINE=9633799; PubMed=8688087;  
 RA Bult C.J., White O., Olsen G.J., Zhou L., Fleischmann R.D.,  
 RA Sutton G.G., Blake J., Fitzgerald L.M., Clayton R.A., Gocayne J.D.,  
 RA Kerecavage A.R., Dougherty B.A., Tomb J.-F., Adams M.D., Reich C.I.,  
 RA Overbeek R., Kirkness E.F., Weinstock K.G., Merrick J.M., Glodek A.,  
 RA Scott J.L., Geoghegan N.S.M., Weidman J.F., Fuhmann J.L., Nguyen D.,  
 RA Uterback T.R., Kelley J.M., Peterson J.D., Sadow P.W., Hanna M.C.,  
 RA Cotton M.D., Roberts K.M., Hurst M.A., Kaine B.P., Borodovsky M.,  
 RA Klenk H.-P., Fraser C.M., Smith H.O., Woese C.R., Venter J.C.;  
 RA "Complete genome sequence of the methanogenic archaeon, Methanococcus  
 RT jannaschii."  
 RT Science 273:1058-1073(1996).  
 RL -1- SIMILARITY: SOME, TO COENZYME F420 HYDROGENASE GAMMA SUBUNIT.  
 CC -1- SIMILARITY: SOME, TO COENZYME F420 HYDROGENASE GAMMA SUBUNIT.  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation-  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL: U67519; AAB98722.1; -  
 DR PIR: F64390; F64390.  
 DR TIGR: MJ0726; -  
 DR InterPro: IPR001450; 4FE4S\_ferredoxin.  
 DR InterPro: IPR006137; Oxidored\_g6.  
 DR Pfam: PF00037; fer4; 1.  
 DR Pfam: PF01058; oxidored\_g6; 1.  
 DR PROSITE: PS00198; 4FE4S\_FERREDOXIN; 1.  
 KW Hypothetical protein; Oxidoreductase; Iron-sulfur; 4Fe-4S;  
 KW Electron transport; Complete proteome.  
 FT METRL 169 169 IRON-SULFUR 1 (4FE-4S) (POTENTIAL).  
 FT METRL 172 172 IRON-SULFUR 1 (4FE-4S) (POTENTIAL).  
 FT METRL 175 175 IRON-SULFUR 1 (4FE-4S) (POTENTIAL).  
 FT METRL 179 179 IRON-SULFUR 2 (4FE-4S) (POTENTIAL).  
 FT METRL 197 197 IRON-SULFUR 2 (4FE-4S) (POTENTIAL).  
 FT METRL 200 200 IRON-SULFUR 2 (4FE-4S) (POTENTIAL).  
 FT METRL 203 203 IRON-SULFUR 2 (4FE-4S) (POTENTIAL).  
 FT METRL 207 207 IRON-SULFUR 1 (4FE-4S) (POTENTIAL).  
 SO SEQUENCE 216 AA; 23789 MW; 2FDP067698982FE CRC64.  
 Query Match 1.5%; Score 8; DB 1; Length 216;  
 Best Local Similarity 100.0%; Pred. No. 4.4;  
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 Oy 101 KEIYDSTL 108  
 Db 117 KEIYDSTL 124  
 RESULT 11  
 BORA\_MOUSE STANDARD; PRT; 349 AA.  
 AC O9JMS6; O9JMS6; 41. Created)  
 DT 28-FEB-2003 (Rel. 41, Last sequence update)  
 DT 10-OCT-2003 (Rel. 42, Last annotation update)  
 DE Cdc42 effector protein 4 (Bander of Rho GTPase 4).  
 GN CDC42EP4 OR BORG4 OR CEP4.  
 OS Mus musculus (Mouse).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 OX NCBI\_TaxID=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A., AND TISSUE SPECIFICITY.  
 RC TISSUE=ileal mucosa;  
 RX MEDLINE=21036164; PubMed=11185749;  
 RA Osada N., Kusuda T., Suzuki Y., Sugano S., Hashimoto K.;  
 RT "Sequence analysis, gene expression, and chromosomal assignment of  
 RL mouse BORG4 gene and its human orthologue."  
 RL J. Hum. Genet. 45:374-377(2000).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC MEDLINE=22388257; PubMed=12477932;  
 RA Straube R.L., Feingold E.A., Grouse L.H., Derge J.G.,  
 RA Klausner R.D., Collins F.S., Wagner L., Shemmen C.M., Schuler G.D.,  
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bat N.K.,  
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,  
 RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,  
 RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,  
 RA Brownstein M.J., Ueda T.B., Toshitsuki S., Carninci P., Prange C.,  
 RA Rana S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullany S.J.,  
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,  
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,  
 RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,  
 RA Whiting M., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A.,  
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,  
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,  
 RA Butlerfield Y.S.N., Krzyzanski M.I., Skalska U., Smalins D.E.,  
 RA Schercher A., Schein J.E., Jones S.J.M., Marra M.A.;  
 RT "Generation and initial analysis of more than 15,000 full-length  
 RT human and mouse cDNA sequences."  
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).  
 RN [3]  
 RP SEQUENCE OF 1-143 FROM N.A., AND INTERACTION WITH TC10 AND CDC42.  
 RX MEDLINE=99421943; PubMed=10490598;  
 RA Joberty G., Perlungher R.R., Macara I.G.;  
 RT "The Borge, a new family of Cdc42 and Tc10 GTPase-interacting  
 RT proteins."  
 RL Mol. Cell. Biol. 19:6585-6597(1999).  
 CC -1- FUNCTION: Probably involved in the organization of the actin  
 CC cytoskeleton. May act downstream of CDC42 to induce actin  
 CC filament assembly leading to cell shape changes. Induces  
 CC pseudopodia formation, when overexpressed in fibroblasts.  
 CC -1- SUBUNIT: Interacts with CDC42 and Tc10, in a GTP-dependent manner.  
 CC -1- SIMILARITY: BELONGS TO THE BORG/CEP FAMILY.  
 CC -1- SIMILARITY: Contains 1 CRIB domain.  
 CC -1- TISSUE SPECIFICITY: Ubiquitous.  
 CC -1- SIMILARITY: BELONGS TO THE BORG/CEP FAMILY.  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation-  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL: AB035088; BAA95932.1; -  
 DR EMBL: BC003857; AA03857.1; -  
 DR EMBL: AF165114; AAD47822.1; -  
 DR MGD: WGI:1929760; Cdc42ep4.  
 DR GO: GO:0005737; C:cytoplasm; IDA.  
 DR GO: GO:0005515; F:protein binding; IPI.  
 DR GO: GO:007266; P:Rho protein signal transduction; IDA.  
 DR InterPro: IPR000095; PAKbox/Rho-binding.  
 DR Pfam: PF00786; PBD; 1.  
 DR PROSITE: PS50108; CRIB; 1.  
 KW Cytoskeleton.  
 FT DOMAIN 27 41 CRIB.  
 FT CONFLICT 143 143 K -> R (IN REF. 3).

Seq	Sequence	349 AA;	37869 MW;	7C44125A7083E16B CRC64;
Query Match				
Best Local Similarity	1.5%;	Score 8;	DB 1;	Length 349;
Matches	8;	Conservative	0;	Mismatches 0;
				Indels 0;
				Gaps 0;
QY	341 SKLPKSL	348		
DB	131 SKLPKSL	138		

RA Hilbert H., Holsappel S., Hosono S., Hullo M.F., Itaya M., Jones L.,  
RA Joris B., Karmata D., Kasahara Y., Klaerr-Blanchard M., Klein C.,  
RA Kobayashi Y., Koetter P., Koningsstein G., Krogh S., Kumano M.,  
RA Kurita K., Lapidus A., Lardino S., Lauder J., Lazarevic V.,  
RA Lee S.M., Levine A., Liu H., Masuda S., Mauel C., Medigone C.,  
RA Medina N., Meliada R.P., Mizuno M., Moestl D., Nakai S., Noback M.,  
RA Noone D., O'Reilly M., Ogawa K., Ogiwara A., Oudega B., Park S.H.,  
RA Parro V., Pohl T.M., Portelle D., Porollis S., Prescott A.M.,  
RA Priesen E., Pujic P., Purnelle B., Rapoport G., Rey M., Reynolds S.,  
RA Rieger M., Rivolet C., Kocha E., Roche B., Rose M., Sadate Y.,  
RA Sato T., Scanlon C., Schleich S., Schroeter R., Scofield F.,  
RA Sekiguchi J., Sekowska A., Seror S.J., Seror P., Shin B.S., Solido B.,  
RA Sorokin A., Tacconi E., Takagi T., Takahashi H., Takemaru K.,  
RA Takeuchi M., Tamkooshi A., Tanaka T., Terpeira P., Tononi A.,  
RA Toesto V., Uchiyama S., Vandenbol M., Vannier P., Vasarotti A.,  
RA Viati A., Wandt R., Wedler E., Wedler H., Weitzenecker T.,  
RA Winters P., Wipat A., Yamamoto H., Yamane K., Yasunoto K., Yata K.,  
RA Yoshida K., Yoshikawa H.F., Zumbstein E., Yoshikawa H., Zumbstein A.,  
RT "The complete genome sequence of the Gram-positive bacterium *Bacillus*  
RT *subtilis*,"  
RT Nature 390:249-256 (1997).  
RL Nature 390:249-256 (1997).  
RN [3]  
RP POSSIBLE FUNCTION.  
RX MEDLINE=2023631; PubMed=10760146;  
RA Que O., Helmann J.D.;  
RT "Manganese homeostasis in *Bacillus subtilis* is regulated by MntR, a  
RT bifunctional regulator related to the diphtheria toxin repressor  
RT family of proteins,"  
RL Mol. Microbiol. 35:1454-1468 (2000).  
CC -1- FUNCTION: This protein is probably a component of a manganese  
CC peremease, a binding protein-dependent, ATP-driven transport  
CC system.  
CC -1- SUBCELLULAR LOCATION: Integral membrane protein (Probable).  
CC -1- SIMILARITY: BELONGS TO THE ABC-3 SUBFAMILY OF INTEGRAL MEMBRANE  
CC PROTEINS.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
CC EMBL, AF008220; AAC00231.1;  
DR EMBL; Z91119; CAB15053.1; -;  
DR PIR; D69992; D69992.  
DR Subtilist; BG13853; mntC.  
DR InterPro; IPR001626; ABC\_transp3.  
DR Pfam; PF00950; ABC-3; 1.  
DR Transport; Ion transport; Transmembrane; Complete proteome.  
FT TRANSMEM 17 37 POTENTIAL.  
FT TRANSMEM 42 62 POTENTIAL.  
FT TRANSMEM 68 88 POTENTIAL.  
FT TRANSMEM 98 118 POTENTIAL.  
FT TRANSMEM 143 163 POTENTIAL.  
FT TRANSMEM 166 186 POTENTIAL.  
FT TRANSMEM 189 209 POTENTIAL.  
FT TRANSMEM 228 248 POTENTIAL.  
FT TRANSMEM 255 275 POTENTIAL.  
SQ SEQUENCE 435 AA; 47944 MW; 68B6D590D90A73BC CRC64;  
Query Match 1.5%; Score 8; DB 1; Length 435;  
Best Local Similarity 100.0%; Pred. No. 8.1;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

ID FA10 RABIT STANDARD; PRT; 490 AA.  
AC O19045;  
DT 15-DEC-1998 (Rel. 37, Created)  
DT 15-DEC-1998 (Rel. 37, Last sequence update)  
DT 10-OCT-2003 (Rel. 42, Last annotation update)  
DE Coagulation factor X precursor (EC 3.4.21.6) (Stuart factor).  
GN F10  
OS *Oryctolagus cuniculus* (Rabbit).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.  
OX NCBI\_Taxid=9986;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=9725631; PubMed=9101642;  
RT Pendurthi U.R., Anderson K.D., James H.L.;  
RT "Characterization of a full-length cDNA for rabbit factor X,"  
RL Thromb. Res. 85:503-514 (1997).  
CC -1- FUNCTION: Factor Xa is a vitamin K-dependent glycoprotein that  
CC converts prothrombin to thrombin in the presence of factor Va,  
CC calcium and phospholipid during blood clotting.  
CC -1- CATALYTIC ACTIVITY: Preferential cleavage: Arg-|-Thr and then  
CC Arg-|-Ile bonds in prothrombin to form thrombin.  
CC -1- SUBUNIT: The two chains are formed from a single-chain precursor  
CC by the excision of two Arg residues and are held together by 1 or  
CC more disulfide bonds.  
CC -1- PTM: The vitamin K-dependent, enzymatic carboxylation of some  
CC glutamate residues allows the modified protein to bind calcium (By  
CC similarity).  
CC -1- PTM: N- and O-glycosylated (By similarity).  
CC -1- PTM: THE ACTIVATION PEPTIDE IS CLEAVED BY FACTOR IXA (IN THE  
CC INTRINSIC PATHWAY), OR BY FACTOR VIIA (IN THE EXTRINSIC PATHWAY)  
CC (BY SIMILARITY).  
CC -1- MISCELLANEOUS: Calcium also binds, with stronger affinity to  
CC another site, beyond the Gla domain.  
CC -1- SIMILARITY: Belongs to peptidase family S1.  
CC -1- SIMILARITY: Contains 2 EGF-like domains.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
CC EMBL, AF003200; AAB62542.1; -;  
DR HSSP; P00742; IHCG.  
DR MEROPS; S01.216; -;  
DR InterPro; IPR000152; Aex\_hydroxyl\_S.  
DR InterPro; IPR009003; Cys\_Ser\_trypsin.  
DR InterPro; IPR000742; EGF\_2.  
DR InterPro; IPR001881; EGF\_Ca.  
DR InterPro; IPR001438; EGF\_II.  
DR InterPro; IPR006209; EGF\_like.  
DR InterPro; IPR002383; Gla\_blood.  
DR InterPro; IPR001254; Peptidase\_S1.  
DR InterPro; IPR001314; Peptidase\_S1A.  
DR InterPro; IPR000294; VitK\_dep\_Gla.  
DR Pfam; PF00008; EGF; 2.  
DR Pfam; PF00594; Gla; 1.  
DR Pfam; PF00089; trypsin; 1.  
DR PRINTS; PR00722; CHYTOTRYPIN.  
DR PRINTS; PR00010; EGFBLD.  
DR PRINTS; PR00001; GLABLOOD.  
DR SMART; SM00179; EGF\_CA; 1.  
DR SMART; SM00069; Gla; 1.  
DR SMART; SM00020; Tryp\_Spc; 1.  
DR PROSITE; PS00010; AEX\_HYDROXYL; 1.  
DR PROSITE; PS00022; EGF\_1; 1.  
DR PROSITE; PS01186; EGF\_2; 2.  
DR PROSITE; PS00026; EGF\_3; 1.  
DR PROSITE; PS01187; EGF\_CA; 1.  
DR PROSITE; PS00011; GLU\_CARBOXYLATION; 1.

```

DR PROSITE; PS50240; TRYPSIN DOM; 1.
DR PROSITE; PS00134; TRYPSIN HIS; 1.
DR PROSITE; PS00135; TRYPSIN SER; 1.
KW Glycoprotein; Hydrolase; Serine protease; Plasma; Blood coagulation;
KW Gamma-carboxyglutamic acid; Hydroxylation; Calcium-binding; Vitamin K;
KW Signal; Zymogen; EGF-like domain; Repeat.
FT SIGNAL 1 20
FT PROPEP 1 40
FT CHAIN 41 180
FT CHAIN 184 490
FT PROPEP 184 490
FT CHAIN 233 490
FT CHAIN 233 490
FT DOMAIN 86 122
FT DOMAIN 125 165
FT DOMAIN 233 490
FT MOD_RES 46 46
FT MOD_RES 47 47
FT MOD_RES 47 47
FT MOD_RES 54 54
FT MOD_RES 54 54
FT MOD_RES 56 56
FT MOD_RES 59 59
FT MOD_RES 60 60
FT MOD_RES 65 65
FT MOD_RES 66 66
FT MOD_RES 69 69
FT MOD_RES 72 72
FT MOD_RES 75 75
FT MOD_RES 79 79
FT MOD_RES 103 103
FT ACT_SITE 274 274
FT ACT_SITE 320 320
FT ACT_SITE 417 417
FT DISULFID 90 101
FT DISULFID 95 110
FT DISULFID 112 121
FT DISULFID 129 140
FT DISULFID 136 149
FT DISULFID 151 164
FT DISULFID 172 340
FT DISULFID 239 244
FT DISULFID 259 275
FT DISULFID 388 402
FT DISULFID 413 441
FT CARBOHYD 61 61
FT CARBOHYD 187 187
FT CARBOHYD 205 205
SQ SEQUENCE 490 AA; 53965 MW; 3A39FA85F2A6D11 CRC64;

Query Match 1.5%; Score 8; DB 1; Length 490;
Best Local Similarity 100.0%; Pred. No. 9;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 4 LGALAG 11
DB 9 LGALAG 16

```

```

DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Hypothetical protein yncp.
GN YHCP OR B3240
OS Escherichia coli.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Escherichia.
OC NCBI_TaxID=562;
RP STRAIN=K12 / MG1655;
RX MEDLINE=9742617; PubMed=9278503;
RA Blaxter F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,
RA Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,
RA Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,
RA Mau B., Shao Y.;
RT "The complete genome sequence of Escherichia coli K-12."
RL Science 277:1453-1474(1997).
CC -1- SUBCELLULAR LOCATION: Integral membrane protein (Potential).
CC -1- SIMILARITY: SOME, TO B.CEPACIA FUSC.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; U18997; AA56042.1; -
DR EMBL; AE000403; AAC76272.1; -
DR PIR; B65116; B65116.
DR Ecogen; EG12822; yncp.
DR InterPro; IPR006726; FUSC.
DR Pfam; PF04632; FUSC. 1.
KW Hypothetical protein; Transmembrane; Complete proteome.
FT TRANSMEM 13 33
FT TRANSMEM 38 58
FT TRANSMEM 69 89
FT TRANSMEM 93 113
FT TRANSMEM 121 141
FT TRANSMEM 152 172
FT TRANSMEM 370 390
FT TRANSMEM 407 427
FT TRANSMEM 431 451
FT TRANSMEM 459 479
FT TRANSMEM 482 502
SQ SEQUENCE 655 AA; 73591 MW; 6C559FC19C276656 CRC64;

Query Match 1.5%; Score 8; DB 1; Length 655;
Best Local Similarity 100.0%; Pred. No. 12;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 74 NATEEIL 81
DB 331 NATEEIL 338

```

Search completed: May 5, 2004, 13:46:13  
Job time : 11.4395 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 5, 2004, 13:41:29 ; Search time 37.7812 Seconds  
(without alignments)  
4325.921 Million cell updates/sec

Title: US-09-743-684a-1\_COPY\_7\_524

Perfect score: 518  
Sequence: 1 LASLGAALAGPVGLKECT.....NTETAAQCNAYEHCKRHWN 518

Scoring table: OLIGO

Gapop 60.0 , Gapext 60.0

Searched: 1017041 seqs, 315518202 residues

Word size : 0

Total number of hits satisfying chosen parameters: 1017041

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Database :

SPTREMBL\_25:\*

- 1: sp\_archaea:\*
- 2: sp\_bacteria:\*
- 3: sp\_fungi:\*
- 4: sp\_human:\*
- 5: sp\_invertebrate:\*
- 6: sp\_mammal:\*
- 7: sp\_mhc:\*
- 8: sp\_organelle:\*
- 9: sp\_phage:\*
- 10: sp\_plant:\*
- 11: sp\_ricket:\*
- 12: sp\_virus:\*
- 13: sp\_vertebrate:\*
- 14: sp\_unclassified:\*
- 15: sp\_virus:\*
- 16: sp\_bacteriophage:\*
- 17: sp\_archaeal:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	22	4.2	554	11	Q8BFP1 mus musculus
2	20	3.9	121	6	P79254 ovis aries
3	13	2.5	512	13	Q7SY70 xenopus lae
4	12	2.3	449	11	Q8BJV5 mus musculus
5	12	2.3	525	11	Q8C1C1 mus musculus
6	11	2.1	402	11	Q8C1M0 mus musculus
7	9	1.7	520	13	Q8UVZ4 brachydanio
8	9	1.7	522	13	Q8D682 brachydanio
9	8	1.5	50	9	Q854G1 mycobacteri
10	8	1.5	177	4	Q9NSM6 homo sapien
11	8	1.5	201	2	Q44889 bordetella
12	8	1.5	201	16	Q7WGY5 bordetella
13	8	1.5	294	13	Q98S10 brachydanio
14	8	1.5	297	16	Q7VPM8 haemophilus
15	8	1.5	300	2	Q9JUN6 vibrio vuln
16	8	1.5	300	16	Q8D492 vibrio vuln

17	8	1.5	326	16	Q8X584 escherichia
18	8	1.5	326	16	Q83RP7 shigella fl
19	8	1.5	330	2	Q9X7G2 methylobact
20	8	1.5	337	12	Q8UT77 lumpy skin
21	8	1.5	337	12	Q8UTN0 lumpy skin
22	8	1.5	337	12	Q91ML5 lumpy skin
23	8	1.5	354	16	Q8YB05 brucella me
24	8	1.5	373	2	Q937R1 brucella me
25	8	1.5	375	16	Q8FMW3 brucella su
26	8	1.5	390	5	Q862X0 plasmodium
27	8	1.5	405	16	Q8XSP7 ralsiconia s
28	8	1.5	456	2	Q93CS1 shigella bo
29	8	1.5	655	16	Q8X9E6 escherichia
30	8	1.5	655	16	Q8PD51 escherichia
31	8	1.5	655	16	Q83JE2 shigella fl
32	8	1.5	663	2	Q841B5 neisseria m
33	8	1.5	775	5	Q81JP7 plasmodium
34	8	1.5	1160	16	Q88D24 pseudomonas
35	7	1.4	13	4	Q75905 homo sapien
36	7	1.4	55	16	Q7VBC3 prochloroxo
37	7	1.4	61	2	Q490E5 mycoplasma
38	7	1.4	68	9	Q856S0 mycobacteri
39	7	1.4	80	16	Q8RC04 thermococ
40	7	1.4	83	2	Q9AED2 lactobacilli
41	7	1.4	90	9	Q9ZXL7 bacterioph
42	7	1.4	94	16	Q9S8S9 streptococ
43	7	1.4	101	17	Q8PMG5 streptococ
44	7	1.4	102	16	Q99XL4 streptococ
45	7	1.4	102	16	Q8E2R0 streptococ

#### ALIGNMENTS

##### RESULT 1

Q8BFP1	PRELIMINARY;	PRT;	554 AA.
AC Q8BFP1	01-MAR-2003 (TREMBLrel. 23, Created)		
DT 01-MAR-2003	(TREMBLrel. 23, Last sequence update)		
DT 01-OCT-2003	(TREMBLrel. 25, Last annotation update)		
DE	Proteasom.		
GN	PSAP.		
OS	Mus musculus (Mouse).		
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.		
OX	NCBI_TaxID=10090;		
RN	[1]		
RP	SEQUENCE FROM N. A.		
RC	STRAIN=NOB; TISSUE=Kidney, and Thymus;		
RX	MEDLINE=22354683; PubMed=12466851;		
RA	The FANTOM Consortium,		
RA	"The RIKEN Genome Exploration Research Group Phase I & II Team;		
RT	"Analysis of the mouse transcriptome based on functional annotation of		
RT	60,770 full-length cDNAs."		
RL	Nature 420:563-573(2002).		
DR	EMBL; AK088369; BAC40308.1; -		
DR	EMBL; AK089998; BAC41035.1; -		
DR	MGD; MGI:97783; Psap.		
DR	GO; GO:0005764; C:lysosome; IEA.		
DR	GO; GO:0006655; P:sphingolipid metabolism; IEA.		
DR	InterPro; IPR003119; SAPA.		
DR	InterPro; IPR007856; SAPB_1.		
DR	InterPro; IPR008138; SAPB_2.		
DR	InterPro; IPR008140; SAPB sub.		
DR	InterPro; IPR008373; Saposin.		
DR	InterPro; IPR008139; Saposinb.		
DR	Pfam; PF02199; SAPA_2.		
DR	Pfam; PF05184; SAPB_1; 4.		
DR	Pfam; PF03489; SAPB_2; 4.		
DR	PRINTS; PR01797; SAPOSIN.		
DR	ProDom; PD001732; SAPB sub; 3.		
DR	SMART; SM00162; SAPA_2.		

DR SMART; SM00118; SAPB; 4.  
 SQ SEQUENCE 554 AA; 61050 MW; FFS8DB79C7CC0C18 CRC64;  
 Query Match  
 Best Local Similarity 4.2%; Score 22; DB 11; Length 554;  
 Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 425 EILAALEKGCSEFLPDYKOCOD 446  
 DB 461 EILAALEKGCSEFLPDYKOCOD 482  
 RESULT 2  
 ID P79254 PRELIMINARY; PRT; 121 AA.  
 AC P79254;  
 DT 01-MAY-1997 (TREMBLrel. 03, Created)  
 DT 01-MAY-1997 (TREMBLrel. 03, Last sequence update)  
 DE 01-OCT-2003 (TREMBLrel. 25, Last annotation update)  
 DE Sulfated glycoprotein-1/SGP-1 (Fragment).  
 OS Ovis aries (Sheep).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;  
 OC Bovidae; Caprinae; Ovis.  
 OX NCBI\_TaxID=9940;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA MEDLINE=96249303; PubMed=8848570;  
 RA Spencer T.E., Graf G.H., Bazer F.W.;  
 RT "Sulfated glycoprotein-1 (SGP-1) expression in ovine endometrium  
 during the oestrous cycle and early pregnancy.";  
 RL Reprod. Fert. 11. Dev. 7:1053-1060(1995).  
 DR EMBL; S82555; AAD1405.1; -.  
 DR GO; GO:0005764; C:lysosome; IEA.  
 DR GO; GO:0006665; P:sphingolipid metabolism; IEA.  
 DR InterPro; IPR008138; SAPB 2.  
 DR InterPro; IPR008139; SaposinB.  
 DR Pfam; PF03489; SAPB\_2; 1.  
 DR PRINTS; PR01797; Saposin.  
 DR SMART; SM00118; SAPB; 1.  
 FT NON TER 1  
 SQ SEQUENCE 121 AA; 13604 MW; 4F0F5A6EB83D0C9A CRC64;  
 Query Match  
 Best Local Similarity 3.9%; Score 20; DB 6; Length 121;  
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 219 HYKEBCDRLGPGMADICKNY 238  
 DB 21 HYKEBCDRLGPGMADICKNY 40  
 RESULT 3  
 ID Q7SY70 PRELIMINARY; PRT; 512 AA.  
 AC Q7SY70;  
 DT 01-OCT-2003 (TREMBLrel. 25, Created)  
 DT 01-OCT-2003 (TREMBLrel. 25, Last sequence update)  
 DT 01-OCT-2003 (TREMBLrel. 25, Last annotation update)  
 DE Hypothetical protein.  
 OS Xenopus laevis (African clawed frog).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae; Pipidae;  
 OC Xenoportidae; Xenopus.  
 OX NCBI\_TaxID=8955;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA TISSUE=Whole;  
 RA MEDLINE=22341132; PubMed=12454917;  
 RA Klein S.L., Strausberg R.L., Wagner L., Pontius J., Clifton S.W.,  
 RA Richardson P.;  
 RT "Genetic and genomic tools for Xenopus research: The NIH Xenopus

RT Initiative.";  
 RL Dev. Dyn. 225:384-391(2002).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Whole;  
 RX MEDLINE=22388257; PubMed=12477932;  
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,  
 RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,  
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,  
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Heien F.,  
 RA Diachenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,  
 RA Stappleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,  
 RA Brownstein M.J., Usdin T.B., Tothiyuki S., Carninci P., Prange C.,  
 RA Rata S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullany S.J.,  
 RA Bosak S.A., McGwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,  
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Huiyk S.W.,  
 RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,  
 RA Fahy J., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A.,  
 RA Whiting U., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,  
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickinson M.C.,  
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,  
 RA Krzywinski M.I., Skalka U., Smalins D.E., Scherch A., Schein J.E.,  
 RA Jones S.J., Marra M.A.;  
 RT "Generation and initial analysis of more than 15,000 full-length human  
 and mouse cDNA sequences.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RA TISSUE=Whole;  
 RC Klein S., Strausberg R.;  
 RA Submitted (JUL-2003) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; BC054988; AAH54988.1; -.  
 DR HYPOHETICAL protein.  
 SQ SEQUENCE 512 AA; 57555 MW; 57CFASR2093FEB1 CRC64;  
 Query Match  
 Best Local Similarity 2.5%; Score 13; DB 13; Length 512;  
 Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 46 VNNKPTVKSLPCD 58  
 DB 52 VNNKPTVKSLPCD 64  
 RESULT 4  
 ID Q8BV5 PRELIMINARY; PRT; 449 AA.  
 AC Q8BV5;  
 DT 01-MAR-2003 (TREMBLrel. 23, Created)  
 DT 01-MAR-2003 (TREMBLrel. 23, Last sequence update)  
 DT 01-OCT-2003 (TREMBLrel. 25, Last annotation update)  
 DE Hypothetical saposin A-type domain/saposin type B containing  
 DE protein.  
 GN 2310020A2RIK.  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 OX NCBI\_TaxID=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=C57BL/6J; TISSUE=Eye;  
 RX MEDLINE=22354683; PubMed=12466851;  
 RA The FANTOM Consortium;  
 RA "The RIKEN Genome Exploration Research Group Phase I & II Team;  
 RT "Analysis of the mouse transcriptome based on functional annotation of  
 RT 60,770 full-length cDNAs.";  
 RL Nature 420:563-573(2002).  
 DR EMBL; AK078699; BAC37363.1; -.  
 DR MGP; MGI:1924193; 2310020A2RIK.  
 DR GO; GO:0005764; C:lysosome; IEA.  
 DR GO; GO:0006665; P:sphingolipid metabolism; IEA.  
 DR InterPro; IPR003119; Sapa.  
 DR InterPro; IPR007856; Sapa\_1.

DR InterPro; IPR008138; SAPB\_2.  
DR InterPro; IPR008140; SAPB\_sub.  
DR InterPro; IPR008373; Saposin.  
DR InterPro; IPR008139; SaposinB.  
DR Pfam; PF02199; SAPA\_1.  
DR Pfam; PF05184; SAPB\_1; 3.  
DR Pfam; PF03489; SAPB\_2; 3.  
DR PRINTS; PR01797; SAPOSIN.  
DR ProDom; PD001732; SAPB\_sub; 2.  
DR SMART; SM00162; SAPA\_1.  
DR SMART; SM00118; SAPB; 3.  
DR Hypothetical protein.  
SQ SEQUENCE 449 AA; 48705 MW; 377CA2A475B292F CRC64;

Query Match 2.3%; Score 12; DB 11; Length 449;  
Best Local Similarity 100.0%; Pred. No. 0.0067;  
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 46 VWNKPTVKSLLPC 57  
DB 56 VWNKPTVKSLLPC 67

## RESULT 5

O8C1C1 PRELIMINARY; PRT; 525 AA.

DT 01-MAR-2003 (TREMBLrel. 23, Created)  
DT 01-MAR-2003 (TREMBLrel. 23, Last sequence update)  
DT 01-OCT-2003 (TREMBLrel. 25, Last annotation update)  
DE Hypothetical saposin A-type domain/saposin type B containing  
DE protein.  
GN 2310020A21RIK.  
OS Mus musculus (Mouse).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
OX NCBI\_TaxID=10090;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=C57BL/6J; TISSUE=Skin;  
RX MEDLINE=22354683; PubMed=12466851;  
RA The FANTOM Consortium;  
RA The RIKEN Genome Exploration Research Group Phase I & II Team;  
RT "Analysis of the mouse transcriptome based on functional annotation of  
RT 60,770 full-length cDNAs."  
RL Nature 420:563-573(2002).  
DR EMBL; AK028455; BAC25961.1; -;  
DR MGD; MGI:1924193; 2310020A21RIK.  
DR GO; GO:0005764; C:lysosome; IEA.  
DR GO; GO:0006665; P:sphingolipid metabolism; IEA.  
DR InterPro; IPR003119; SAPA.  
DR InterPro; IPR007856; SAPB\_1.  
DR InterPro; IPR008138; SAPB\_2.  
DR InterPro; IPR008140; SAPB\_sub.  
DR InterPro; IPR008373; Saposin.  
DR InterPro; IPR008139; SaposinB.  
DR Pfam; PF02199; SAPA\_1; 2.  
DR Pfam; PF05184; SAPB\_1; 4.  
DR Pfam; PF03489; SAPB\_2; 4.  
DR PRINTS; PR01797; SAPOSIN.  
DR ProDom; PD001732; SAPB\_sub; 3.  
DR SMART; SM00162; SAPA\_1; 2.  
DR SMART; SM00118; SAPB; 3.  
DR Hypothetical protein.  
SQ SEQUENCE 525 AA; 57350 MW; 374F6050CDC4D223 CRC64;

Query Match 2.3%; Score 12; DB 11; Length 525;  
Best Local Similarity 100.0%; Pred. No. 0.0076;  
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 46 VWNKPTVKSLLPC 57  
DB 56 VWNKPTVKSLLPC 67

RESULT 6  
O8C1N0 PRELIMINARY; PRT; 402 AA.

ID O8C1N0  
AC O8C1N0  
DT 01-MAR-2003 (TREMBLrel. 23, Created)  
DT 01-MAR-2003 (TREMBLrel. 23, Last sequence update)  
DT 01-OCT-2003 (TREMBLrel. 25, Last annotation update)  
DE Hypothetical saposin A-type domain/saposin type B containing  
DE protein.  
GN 2310020A21RIK.  
OS Mus musculus (Mouse).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
OX NCBI\_TaxID=10090;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=C57BL/6J; TISSUE=Tongue;  
RX MEDLINE=22354683; PubMed=12466851;  
RA The FANTOM Consortium;  
RA The RIKEN Genome Exploration Research Group Phase I & II Team;  
RT "Analysis of the mouse transcriptome based on functional annotation of  
RT 60,770 full-length cDNAs."  
RL Nature 420:563-573(2002).  
DR EMBL; AK009408; BAC25258.1; -;  
DR MGD; MGI:1924193; 2310020A21RIK.  
DR GO; GO:0005764; C:lysosome; IEA.  
DR GO; GO:0006665; P:sphingolipid metabolism; IEA.  
DR InterPro; IPR003119; SAPA.  
DR InterPro; IPR007856; SAPB\_1.  
DR InterPro; IPR008138; SAPB\_2.  
DR InterPro; IPR008140; SAPB\_sub.  
DR InterPro; IPR008373; Saposin.  
DR InterPro; IPR008139; SaposinB.  
DR Pfam; PF02199; SAPA\_1; 1.  
DR Pfam; PF05184; SAPB\_1; 3.  
DR Pfam; PF03489; SAPB\_2; 3.  
DR PRINTS; PR01797; SAPOSIN.  
DR ProDom; PD001732; SAPB\_sub; 2.  
DR SMART; SM00162; SAPA\_1; 1.  
DR SMART; SM00118; SAPB; 3.  
DR Hypothetical protein.  
SQ SEQUENCE 402 AA; 44420 MW; E90017CBF4017ED6 CRC64;

Query Match 2.1%; Score 11; DB 11; Length 402;  
Best Local Similarity 100.0%; Pred. No. 0.06;  
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 448 FVAEYEPVLI 458  
DB 322 FVAEYEPVLI 332

## RESULT 7

O8UVZ4 PRELIMINARY; PRT; 520 AA.

ID O8UVZ4  
AC O8UVZ4  
DT 01-MAR-2002 (TREMBLrel. 20, Created)  
DT 01-MAR-2002 (TREMBLrel. 20, Last sequence update)  
DT 01-OCT-2003 (TREMBLrel. 25, Last annotation update)  
DE Lysosomal cofactor/neurotrophic factor prosaposin.  
OS Brachyotia reitio (Zebrafish) (Danio rerio).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;  
OC Cyprinidae; Danio.  
OX NCBI\_TaxID=7955;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC Seo H.-C., Lie O., Fjose A., O'Brien J.S., Kishimoto Y.,  
RA "Cloning, expression and promoter analysis of zebrafish prosaposin."  
RT Submitted (JUN-2000) to the EMBL/Genbank/DBJ databases.  
RL EMBL; AF276996; AAL54381.1; -;

DR GO:0005764; C:lysosome; IEA.  
 DR GO:0006655; P:sphingolipid metabolism; IEA.  
 DR InterPro: IPR003119; SApA.  
 DR InterPro: IPR007856; SApB\_1.  
 DR InterPro: IPR008138; SApB\_2.  
 DR InterPro: IPR008140; SApB\_sub.  
 DR InterPro: IPR008373; Saposin.  
 DR InterPro: IPR008139; SaposinB.  
 DR Pfam: PF02199; SApA; 2.  
 DR Pfam: PF05184; SApB\_1; 3.  
 DR Pfam: PF03489; SApB\_2; 4.  
 DR PRINTS: PR01797; SApOSIN.  
 DR PRODOM: PD001732; SApB\_sub; 3.  
 DR SMART: SM00162; SApA; 2.  
 DR SMART: SM00118; SApB; 4.  
 SQ SEQUENCE 520 AA; 57431 MW; F9E620F84BA1CB5 CRC64;

Query Match 1.7%; Score 9; DB 13; Length 520;  
 Best Local Similarity 100.0%; Pred. No. 7.4;  
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 26 WCONVKTAS 34  
 DB 30 WCONVKTAS 38

RESULT 8  
 ID Q9DG82 PRELIMINARY; PRT; 522 AA.  
 AC Q9DG82;  
 DT 01-MAR-2001 (TrEMBLrel. 16, Created)  
 DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)  
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
 DE Prosaposin.  
 GN PSAP.  
 OS Brachydanio rerio (Zebrafish) (Danio rerio).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;  
 OC Cyprinidae; Danio.  
 OC NCBI\_TaxId=7955;  
 OX NCBI\_TaxId=7955;  
 RN  
 RP SEQUENCE FROM N.A.  
 RA Altman N., Horowitz M.;  
 RT "The zebrafish prosaposin cDNA."  
 RT Submitted (NOV-1998) to the EMBL/GenBank/DBJ databases.  
 DR EMBL: AF108655: AAG32919.1; --  
 DR ZFIN: ZDB-GENE-020108-1; psap.  
 DR GO:0005764; C:lysosome; IEA.  
 DR GO:0006655; P:sphingolipid metabolism; IEA.  
 DR InterPro: IPR003119; SApA.  
 DR InterPro: IPR007856; SApB\_1.  
 DR InterPro: IPR008138; SApB\_2.  
 DR InterPro: IPR008140; SApB\_sub.  
 DR InterPro: IPR008373; Saposin.  
 DR InterPro: IPR008139; SaposinB.  
 DR Pfam: PF02199; SApA; 2.  
 DR Pfam: PF05184; SApB\_1; 3.  
 DR Pfam: PF03489; SApB\_2; 4.  
 DR PRINTS: PR01797; SApOSIN.  
 DR PRODOM: PD001732; SApB\_sub; 3.  
 DR SMART: SM00162; SApA; 2.  
 DR SMART: SM00118; SApB; 4.  
 SQ SEQUENCE 522 AA; 57671 MW; D3C15A305725C1CD CRC64;

Query Match 1.7%; Score 9; DB 13; Length 522;  
 Best Local Similarity 100.0%; Pred. No. 7.4;  
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 26 WCONVKTAS 34  
 DB 30 WCONVKTAS 38

RESULT 9  
 ID O854G1 PRELIMINARY; PRT; 50 AA.  
 AC O854G1;  
 DT 01-JUN-2003 (TrEMBLrel. 24, Created)  
 DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)  
 DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)  
 DE Gp106.  
 GN 106.  
 OS Mycobacteriophage Omega.  
 OC Viruses; dsDNA viruses, no RNA stage; Caudovirales; Siphoviridae.  
 OC NCBI\_TaxId=205879;  
 OX NCBI\_TaxId=205879;  
 RN  
 RP SEQUENCE FROM N.A.  
 RA MEDLINE=22592660; PubMed=12705866;  
 RA Pedulla M.L., Ford M.E., Houtz J.M., Karthikeyan T., Wadsworth C.,  
 RA Lewis J.A., Jacobs-Sera D., Falbo J., Gross J., Panunzio N.R.,  
 RA Brucker W., Kumar V., Kandasamy J., Keenan L., Bardarov S.,  
 RA Krizakov J., Lawrence J.G., Jacobs W.R. Jr., Hendrix R.W.,  
 RA Hatfull G.F.;  
 RT "Origins of highly mosaic mycobacteriophage genomes."  
 RT Cell 113:171-182(2003).  
 RL EMBL: AY129338; AAN12747.1;  
 DR EMBL: AY129338; AAN12747.1;  
 SQ SEQUENCE 50 AA; 5689 MW; 16993D01C4E247F4 CRC64;

Query Match 1.5%; Score 8; DB 9; Length 50;  
 Best Local Similarity 100.0%; Pred. No. 10;  
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 111 ILDIKGE 118  
 DB 43 ILDIKGE 50

RESULT 10  
 ID Q9NSM6 PRELIMINARY; PRT; 177 AA.  
 AC Q9NSM6;  
 DT 01-OCT-2000 (TrEMBLrel. 15, Created)  
 DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)  
 DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)  
 DE Hypothetical protein (Fragment).  
 GN DKFZ761P19121.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 OC NCBI\_TaxId=9606;  
 OX NCBI\_TaxId=9606;  
 RN  
 RP SEQUENCE FROM N.A.  
 RA TISSUE=Amalgam;  
 RA Pousetka A., Melienreuther R., Mewes H.W., Weil B., Wiemann S.;  
 RT Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.  
 DR EMBL: AL162013; CAB82365.1; --  
 DR PIR: T47133; T47133.  
 KW Hypothetical protein.  
 FT NON\_TER  
 FT 1  
 SQ SEQUENCE 177 AA; 20522 MW; EB717DE9993AF417 CRC64;

Query Match 1.5%; Score 8; DB 4; Length 177;  
 Best Local Similarity 100.0%; Pred. No. 29;  
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 425 ETLALEX 432  
 DB 146 ETLALEX 153

RESULT 11  
 ID Q44889 PRELIMINARY; PRT; 201 AA.  
 AC Q44889;  
 DT 01-NOV-1996 (TrEMBLrel. 01, Created)  
 DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)



DT 01-OCT-2003 (TRENBLrel. 25, Last annotation update)  
 DE Fimbrial subunit precursor.  
 GN FIMX.  
 OS Bordetella bronchiseptica (Alcaligenes bronchisepticus).  
 OC Bacteria; Proteobacteria; Betaproteobacteria; Burkholderiales;  
 CC Alcaligenaceae; Bordetella.  
 OX NCBI\_TaxID=518;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=No. 401;  
 RA Savelkoul P., De Keif D., de Groot L., Willems R., Mooi F.,  
 RA van der Zeijst B., Gaastera W.;  
 RT "Characterization of fimb, a gene encoding a fimbrial subunit protein  
 of Bordetella bronchiseptica."  
 RL Submitted (FEB-1993) to the EMBL/Genbank/DBJ databases.  
 DR EMBL: X74118; CA52215.1; -.  
 DR PIR: S36451; S36451.  
 DR GO: GO:0009289; C:fimbria; IEA.  
 DR GO: GO:0007155; P:cell adhesion; IEA.  
 DR InterPro: IPR008966; Adhes\_bact.  
 DR InterPro: IPR000259; Fimbrial.  
 DR Pfam: PF00419; Fimbrial; 1.  
 DR Signal.  
 KW Signal.  
 FT SIGNAL.  
 SQ SEQUENCE 201 AA; 21462 MW; F47670CC05B03F6 CRC64;

Query Match 1.5%; Score 8; DB 2; Length 201;  
 Best Local Similarity 100.0%; Pred. No. 33;  
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 4 LGGALAG 11  
 |||||  
 Db 7 LGGALAG 14

RESULT 12  
 O7MGYS PRELIMINARY; PRT; 201 AA.  
 AC O7MGYS.  
 DT 01-OCT-2003 (TRENBLrel. 25, Created)  
 DT 01-OCT-2003 (TRENBLrel. 25, Last sequence update)  
 DT 01-OCT-2003 (TRENBLrel. 25, Last annotation update)  
 DE Fimbrial protein.  
 GN FIMX OR B3426.  
 OS Bordetella bronchiseptica (Alcaligenes bronchisepticus).  
 OC Bacteria; Proteobacteria; Betaproteobacteria; Burkholderiales;  
 CC Alcaligenaceae; Bordetella.  
 OX NCBI\_TaxID=518;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=RB50 / ATCC BAA-588;  
 RX MEDLINE=22827954; PubMed=12910271;  
 RA Parkhill J., Sebahia M., Preston A., Murphy L.D., Thomson N.,  
 RA Harris D.E., Holden M.T.G., Churcher C.M., Bentley S.D., Mungall K.L.,  
 RA Cerdano-Tarraga A.M., Temple L., James K., Harris B., Quail M.A.,  
 RA Achtman M., Atkin R., Baker S., Basham D., Bason N., Cherevach I.,  
 RA Chillingworth T., Collins M., Cronin A., Davis P., Doggett J.,  
 RA Feltham T., Goble A., Hamlin N., Hauser H., Holroyd S., Jagels K.,  
 RA Leather S., Moule S., Norberczak H., O'Neill S., Ormond D., Price C.,  
 RA Rabinowitsch E., Rutter S., Sanders M., Saunders D., Seeger K.,  
 RA Sharp S., Simmonds M., Skelton J., Squares R., Squares S., Stevens K.,  
 RA Unwin L., Whitehead S., Barrett B.G., Maskell D.J.;  
 RT "Comparative analysis of the genome sequences of Bordetella pertussis,  
 RT Bordetella parapertussis and Bordetella bronchiseptica."  
 RL Nat. Genet. 35:32-40(2003).  
 DR EMBL: BX640447; CAE33918.1; -.  
 KW Complete proteome.  
 SQ SEQUENCE 201 AA; 21511 MW; CB4970CC05439082 CRC64;

Query Match 1.5%; Score 8; DB 16; Length 201;  
 Best Local Similarity 100.0%; Pred. No. 33;  
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 4 LGGALAG 11  
 |||||  
 Db 7 LGGALAG 14

RESULT 13  
 O98S10 PRELIMINARY; PRT; 294 AA.  
 ID O98S10  
 AC O98S10.  
 DT 01-JUN-2001 (TRENBLrel. 17, Created)  
 DT 01-JUN-2001 (TRENBLrel. 17, Last sequence update)  
 DT 01-JUN-2003 (TRENBLrel. 24, Last annotation update)  
 DE Hoxala protein.  
 GN HOXALA.  
 OS Brachydanio rerio (zebrafish) (Danio rerio).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;  
 CC Cyprinidae; Danio.  
 OX NCBI\_TaxID=7955;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA McClinton J.M.;  
 RT "Consequences of Hox gene duplication in the vertebrates: an  
 RT investigation of the zebrafish hox paralogue group 1 genes."  
 RL Submitted (MAR-2001) to the EMBL/Genbank/DBJ databases.  
 CC -1- SUBCELLULAR LOCATION: NUCLEAR (BY SIMILARITY).  
 DR EMBL: AJ306431; CAC34566.1; -.  
 DR HSSP: P14653; 1B72.  
 DR ZFIN: ZDB-GENE-000823-5; hoxala.  
 DR GO: GO:0005634; C:nucleus; IEA.  
 DR GO: GO:0003700; F:transcription factor activity; IEA.  
 DR GO: GO:0006355; P:regulation of transcription, DNA-dependent; IEA.  
 DR InterPro: IPR001356; Homeobox.  
 DR InterPro: IPR000047; HTH\_LambdaRepressor.  
 DR Pfam: PF00046; homeobox; 1.  
 DR PRINTS: PR00024; HOMEBOX.  
 DR PRINTS: PR00031; HTHREPRESSR.  
 DR ProDom: PD000010; Homeobox; 1.  
 DR SMART: SM00389; HOX; 1.  
 DR PROSITE: PS00027; HOMEBOX\_1; 1.  
 DR PROSITE: PS50071; HOMEBOX\_2; 1.  
 KW DNA-binding; Homeobox; Nuclear protein.  
 SQ SEQUENCE 294 AA; 32179 MW; 8A7C2391BF410727 CRC64;

Query Match 1.5%; Score 8; DB 13; Length 294;  
 Best Local Similarity 100.0%; Pred. No. 45;  
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 343 LPKSLSEE 350  
 |||||  
 Db 254 LPKSLSEE 261

RESULT 14  
 O7VPN8 PRELIMINARY; PRT; 297 AA.  
 ID O7VPN8  
 AC O7VPN8.  
 DT 01-OCT-2003 (TRENBLrel. 25, Created)  
 DT 01-OCT-2003 (TRENBLrel. 25, Last sequence update)  
 DT 01-OCT-2003 (TRENBLrel. 25, Last annotation update)  
 DE XerD OR HD0013.  
 GN XERD OR HD0013.  
 OS Haemophilus ducreyi.  
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Pasteurellales;  
 CC Pasteurellaceae; Haemophilus.  
 OX NCBI\_TaxID=730;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=3500HP / ATCC 700724;  
 RA Munson R.S., Jr., Ray W.C., Mahairas G., Sabo P., Mungur R.,  
 RA Johnson L., Nguyen D., Wang J., Forst C., Hood L.;  
 RT "The complete genome sequence of Haemophilus ducreyi."  
 RL Submitted (JUN-2003) to the EMBL/Genbank/DBJ databases.

DR EMBL; AE017151; AAP95036.1; -  
 KW Complete proteome.  
 SQ SEQUENCE 297 AA; 34472 MW; 589A1B3984E73951 CRC64;

Query Match  
 Best Local Similarity 1.5%; Score 8; DB 16; Length 297;  
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 343 LPKSLSEE 350  
 |||||  
 DB 108 LPKSLSEE 115

## RESULT 15

Q93UN6 PRELIMINARY; PRT; 300 AA.  
 ID Q93UN6  
 AC Q93UN6;  
 DT 01-DEC-2001 (TrEMBLrel. 19, Created)  
 DT 01-DEC-2001 (TrEMBLrel. 19, last sequence update)  
 DT 01-OCT-2003 (TrEMBLrel. 25, last annotation update)  
 DE Hypothetical protein.  
 OS Vibrio vulnificus.  
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Vibrionales;  
 OC Vibrionaceae; Vibrrio.  
 OX NCBI\_TaxID=672;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=CM-1;  
 RA Chang M.C., Chang S.Y., Chang C.C.;  
 RT "Partial DNA sequence of the chromosome in Vibrio vulnificus."  
 RL Submitted (MAY-2001) to the EMBL/GenBank/DDJ databases.  
 CC -1 SIMILARITY: BELONGS TO THE LYSR FAMILY OF TRANSCRIPTIONAL  
 CC REGULATORS.  
 DR EMBL; AF080431; AAK50454.1; -  
 DR GO; GO:0003700; F:transcription factor activity; IEA.  
 DR GO; GO:0006355; P:regulation of transcription; DNA-dependent; IEA.  
 DR GO; GO:0006350; P:transcription; IEA.  
 DR InterPro; IPR000847; HTH\_LYR.  
 DR InterPro; IPR005119; LysR\_subst.  
 DR Pfam; PF00126; HTH\_1; 1.  
 DR Pfam; PF03466; LysR\_substrate; 1.  
 KW Hypothetical protein; DNA-binding; Transcription;  
 KW Transcription regulation.  
 SQ SEQUENCE 300 AA; 33710 MW; B7B508680F6D8ED CRC64;

Query Match  
 Best Local Similarity 1.5%; Score 8; DB 2; Length 300;  
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 LIGALAG 11  
 |||||  
 DB 223 LIGALAG 230

Search completed: May 5, 2004, 13:47:42  
 Job time : 38.7812 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 5, 2004, 13:15:48 ; Search time 72.3134 Seconds  
(without alignments)  
2023.963 Million cell updates/sec

Title: US-09-743-684A-1\_COPY\_7\_524  
2759  
Sequence: 1 LASLIGALAGPVLGKRECT.....NTEFAAGCNAYEHCKRHVYN 518

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 1586107 seqs, 282547505 residues  
Total number of hits satisfying chosen parameters: 1586107

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : A\_Geneseq\_29Jan04:\*

1: geneseqp1980s:\*  
2: geneseqp1990s:\*  
3: geneseqp2000s:\*  
4: geneseqp2001s:\*  
5: geneseqp2002s:\*  
6: geneseqp2003as:\*  
7: geneseqp2003bs:\*  
8: geneseqp2004s:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2759	100.0	524	2	AAR70783 Prosapoin
2	2759	100.0	524	2	AAW85652 Human pro
3	2759	100.0	524	3	AAV58716 Human pro
4	2759	100.0	524	6	ABU79099 Lip-TAA b
5	2759	100.0	524	6	ABU05200 Human exp
6	2759	100.0	524	6	ABU05207 Human exp
7	2759	100.0	524	6	ABU05203 Human exp
8	2759	100.0	524	6	ABU07340 Human exp
9	2759	100.0	524	6	ABU05216 Human exp
10	2759	100.0	524	6	ABU05202 Human exp
11	2759	100.0	524	6	ABU05214 Human exp
12	2759	100.0	524	6	ABU05215 Human exp
13	2759	100.0	524	6	ABU05199 Human exp
14	2759	100.0	524	6	ABU05212 Human exp
15	2759	100.0	524	6	ABU05213 Human exp
16	2759	100.0	524	6	ABU05205 Human exp
17	2749	99.6	524	6	ABU05208 Human exp
18	2747.5	99.6	527	4	AAAB31915 Amino aci
19	2747.5	99.6	527	5	ABE68602 Human pan
20	2747.5	99.6	527	6	ABU79100 Lip-TAA b
21	2747.5	99.6	527	6	ABU05204 Human exp
22	2747.5	99.6	527	6	ABU05210 Human exp
23	2742.5	99.4	523	4	AAAB31916 Amino aci
24	2742.5	99.4	523	6	ABU05211 Human exp
25	2738	99.2	526	6	ABU05209 Human exp

26	2737.5	99.2	527	6	ABU05206 Human exp
27	2419.5	87.7	479	6	ABR39442 Human GEN
28	2024.5	73.4	385	6	ABR41750 Human DIT
29	1947	70.6	554	7	ADB85295 Rat tubul
30	1899	68.8	554	5	ABD57102 Mouse lsc
31	1189	43.1	521	3	AAV56966 Human SBP
32	1189	43.1	521	4	AAW78587 Human pro
33	1189	43.1	521	4	AAW79571 Human pro
34	1189	43.1	521	4	AAW86362 Human pro
35	1189	43.1	521	5	AAE15547 Human sec
36	1184	42.9	521	4	AAW79572 Human pro
37	1127.5	40.9	531	4	AAE01770 Human gen
38	1127.5	40.9	531	5	ABG64156 Human alb
39	1125	40.8	210	6	ABU05201 Human exp
40	1093	39.6	209	5	ABG70166 Human pre
41	822.5	29.8	362	4	AAW78588 Human pro
42	822.5	29.8	362	5	ABW07499 Human lip
43	812	29.4	153	6	ABU70504 Human adi
44	811	29.4	153	6	ABU70799 Human adi
45	676	24.5	129	6	ABU70422 Human adi

## ALIGNMENTS

RESULT 1	
AA70783	
ID	AA70783 standard; protein; 524 AA.
AC	AA70783;
XX	
DT	25-MAR-2003 (revised)
DT	30-AUG-1995 (first entry)
XX	
DE	Prosapoin.
XX	
KW	Sapoin-C; neuron; myelination; nervous system; neuroblastoma;
KW	neurotrophic peptide; multiple sclerosis; leukoencephalitis;
KW	adrenal leukodystrophy; prosapoin.
OS	Homo sapiens.
XX	
PN	W09503821-A1.
XX	
PD	09-FEB-1995.
XX	
PF	28-JUL-1994; 94WO-US008453.
XX	
PR	30-JUL-1993; 93US-00100247.
PR	21-APR-1994; 94US-00232513.
PA	(OBRI) O'BRIEN J S.
XX	
PI	O'brien JS, Kishimoto Y,
DR	MPL, 1995-082029/11.
DR	N-PSDB; AAQ85355.
XX	
PT	Stimulating neural cell out-growth and myelination - with prosapoin,
PT	sapoin C or new neurotrophic peptide(s) from cytokine(s), for treating
XX	nervous system diseases.
PS	Disclosure; Page 30-32; 50pp; English.
XX	
CC	The peptide given in AAR70773, corresponding to amino acids 8-29 of human
CC	sapoin-C (AAR70784), promotes neurite outgrowth in vitro. A consensus
CC	sequence was determined by comparing the peptide with hematopoietic and
CC	neuroepithelial cytokines, and neurotrophic peptides (AAR70774-82) were
CC	identified in the AB loop of human ciliary neurotrophic factor,
CC	interleukins-6, -2, -3 and -gamma, erythropoietin and leukocyte
CC	inhibitory factor, and in helix C of human interleukin-1-beta and
CC	oncostatin-M. Prosapoin (AAR70783) and sapoin-C also promoted nerve
CC	cell myelination ex vivo. (Updated on 25-MAR-2003 to correct PN field.)

CC (Updated on 25-MAR-2003 to correct PA field.) (Updated on 25-MAR-2003 to  
 CC correct PI field.)  
 CC  
 XX Sequence 524 AA;

Query Match 100.0%; Score 2759; DB 2; Length 524;  
 Best Local Similarity 100.0%; Pred. No. 7, 2e-235;  
 Matches 518; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LASLLGALAGPVGLGKCTRGSAVWCQNVKTASDCGAVHCLQTVWNNKPTVKSLLPCDIC 60  
 DB 7 LASLLGALAGPVGLGKCTRGSAVWCQNVKTASDCGAVHCLQTVWNNKPTVKSLLPCDIC 66  
 QY 61 KDVTAAADMVKDNATEEELIVLEKTCMDLPKNNASCKEIVDSYLPVILDIKGEYS 120  
 DB 67 KDVTAAADMVKDNATEEELIVLEKTCMDLPKNNASCKEIVDSYLPVILDIKGEYS 126  
 QY 121 RRGVCSALNLCESLQKHLAEINHQLESNKIPELDMTEVVAFPVANIPLLLYPQDGR 180  
 DB 127 RRGVCSALNLCESLQKHLAEINHQLESNKIPELDMTEVVAFPVANIPLLLYPQDGR 186  
 QY 181 SKPQKNDGVCDQCIQMTVDIQTAVRTNSTFVQALVEHKECDRLGPGMADICKNYIS 240  
 DB 187 SKPQKNDGVCDQCIQMTVDIQTAVRTNSTFVQALVEHKECDRLGPGMADICKNYIS 246  
 QY 241 QYSEIATQMMHMQPKICAVGFCDEVKEMPQTLVPAKVASKNVLPALTEVPIKXHE 300  
 DB 247 QYSEIATQMMHMQPKICAVGFCDEVKEMPQTLVPAKVASKNVLPALTEVPIKXHE 306  
 QY 301 VAKSVNVECEVEFLVKEVTKLIDNNKTEKELIDAFDKKCSLPRSLSEECQEVVDITYGS 360  
 DB 307 VAKSVNVECEVEFLVKEVTKLIDNNKTEKELIDAFDKKCSLPRSLSEECQEVVDITYGS 366  
 QY 361 SILSLILEBVEPVLCSMLHCSGTRLPALTVAHVQPKDGGCEVCKLVGYLDRNLEKN 420  
 DB 367 SILSLILEBVEPVLCSMLHCSGTRLPALTVAHVQPKDGGCEVCKLVGYLDRNLEKN 426  
 QY 421 STKQKIIAALKEGSGFLPDPYQKCDQFVAEYEPVLIBEIVMDPSFVCLIKAGCSAH 480  
 DB 427 STKQKIIAALKEGSGFLPDPYQKCDQFVAEYEPVLIBEIVMDPSFVCLIKAGCSAH 486  
 QY 481 KPLGTEKICIMPSYWCQNTERRAACQNAVEHCKHWN 518  
 DB 487 KPLGTEKICIMPSYWCQNTERRAACQNAVEHCKHWN 524

RESULT 2  
 AAW85652  
 ID AAW85652 standard; protein; 524 AA.  
 XX  
 AC AAW85652;  
 XX  
 DT 19-JUL-1999 (first entry)  
 XX  
 DE Human prosaposin N-terminal peptide.  
 XX  
 KM Prosaposin; saposin; prosaptilides; prosaposin receptor agonists; PRA;  
 KM peripheral nervous system; central nervous system; PNS; CNS; Akt; Bcl-2;  
 KM therapy; treatment; apoptosis; caspase; tumour necrosis factor; TNF;  
 KM cytokine; interleukin gamma; IFN; inflammation; rheumatoid arthritis;  
 KM Crohn's disease; irritable bowel syndrome; ashma; cardiac infarction;  
 KM congestive heart failure; multiple sclerosis;  
 KM acute disseminated inflammatory leukoencephalitis;  
 KM progressive multifocal leukoencephalitis;  
 KM Parkinson's disease; amyotrophic lateral sclerosis; Huntington's disease;  
 KM ischemic heart disease; Guillain-Barre disease; alopecia; AIDS dementia;  
 KM cerebral malaria; HTLV; neuropathy;  
 KM inflammatory neurodegenerative disease; toxin-induced liver disease.  
 XX  
 OS Homo sapiens.  
 XX  
 XX  
 PN MO9912559-A1.  
 XX

PD 18-MAR-1999.  
 XX  
 XX  
 PF 09-SEP-1998; 98WO-US019216.  
 PF  
 PR 09-SEP-1997; 97US-0058352P.  
 PR 04-JUN-1998; 98US-0088129P.  
 XX  
 XX  
 PA (REGC) UNIV CALIFORNIA.  
 XX  
 PI O'Brien US;  
 XX  
 DR WPI; 1999-229139/19.  
 DR N-PSDB; AAX08488.  
 XX

Use of prosaposin receptor agonist.

Claim 7; Fig 2; 90pp; English.

Prosaposin is a 70kDa glycoprotein which is proteolytically processed to generate saposins A, B, C and D, all of which are similar to each other and have a similar placement of six cysteines, a glycosylation site and conserved proline residues. Prosaposin, saposin C and prosaposin derived peptides (prosaptilides), have therapeutic applications in promoting recovery after toxic, traumatic, myocardial ischaemic, degenerative and inherited lesions to the peripheral and central nervous system. Prosaposin receptor agonists (PRAs) inhibit proinflammatory cytokine-induced apoptosis by activation of the Ser/Thr protein kinase Akt. Akt releases Bcl-2 and its family members, such as BAD-Bcl-2, which inhibit apoptosis. An additional mechanism which inhibit caspases, thereby apoptosis is by blocking activation of JNK, a proapoptotic signaling component. Within several minutes after binding to the receptor, PRAs block JNK activation induced by tumor necrosis factor-alpha (TNF alpha). The activation of JNK by TNF alpha is another well known mechanism for apoptosis. The method can be used for inhibiting apoptosis which is caspase-mediated or induced by a proinflammatory cytokine which is TNF alpha or interferon-gamma. It can be used for inhibiting apoptosis associated with a disorder such as e.g. rheumatoid arthritis, Crohn's disease, irritable bowel syndrome, asthma, cardiac infarction, congestive heart failure, multiple sclerosis, acute disseminated inflammatory leukoencephalitis, progressive multifocal leukoencephalitis, Alzheimer's disease, Parkinson's disease, amyotrophic lateral sclerosis, Huntington's disease, ischemic heart disease, Guillain-Barre disease, traumatic brain injury, traumatic spinal cord injury, alopecia, AIDS dementia, cerebral malaria, HTLV, neuropathy, inflammatory neurodegenerative disease, and toxin-induced liver disease. This 524 N-terminal peptide of prosaposin also acts as a prosaposin receptor agonist

Sequence 524 AA;

Query Match 100.0%; Score 2759; DB 2; Length 524;  
 Best Local Similarity 100.0%; Pred. No. 7, 2e-235;  
 Matches 518; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LASLLGALAGPVGLGKCTRGSAVWCQNVKTASDCGAVHCLQTVWNNKPTVKSLLPCDIC 60  
 DB 7 LASLLGALAGPVGLGKCTRGSAVWCQNVKTASDCGAVHCLQTVWNNKPTVKSLLPCDIC 66  
 QY 61 KDVTAAADMVKDNATEEELIVLEKTCMDLPKNNASCKEIVDSYLPVILDIKGEYS 120  
 DB 67 KDVTAAADMVKDNATEEELIVLEKTCMDLPKNNASCKEIVDSYLPVILDIKGEYS 126  
 QY 121 RRGVCSALNLCESLQKHLAEINHQLESNKIPELDMTEVVAFPVANIPLLLYPQDGR 180  
 DB 127 RRGVCSALNLCESLQKHLAEINHQLESNKIPELDMTEVVAFPVANIPLLLYPQDGR 186  
 QY 181 SKPQKNDGVCDQCIQMTVDIQTAVRTNSTFVQALVEHKECDRLGPGMADICKNYIS 240  
 DB 187 SKPQKNDGVCDQCIQMTVDIQTAVRTNSTFVQALVEHKECDRLGPGMADICKNYIS 246  
 QY 241 QYSEIATQMMHMQPKICAVGFCDEVKEMPQTLVPAKVASKNVLPALTEVPIKXHE 300

Db 247 QYSEIATOMMHMPKEICALVGFCDVEKEMPMOTLVPKASKNVLPALBELVEPIKHE 306  
 QY 301 VPAKSDVYCEVCEFLVKEVTKLIDNNKTEKEILDAFDMKCSKLPSLSSECOEVDVTGS 360  
 Db 307 VPASDVYCEVCEFLVKEVTKLIDNNKTEKEILDAFDMKCSKLPSLSSECOEVDVTGS 366  
 QY 361 SILSILAEVSPBELVCSMLHLCSGTRLPALTVHTVTPQPDGFCFCEVCKLVGLDRLNLEKN 420  
 Db 367 SILSILAEVSPBELVCSMLHLCSGTRLPALTVHTVTPQPDGFCFCEVCKLVGLDRLNLEKN 426  
 QY 421 STKOEIIAALAEKGSFLPDPYQKODQFVAEYEPVLIEILVWMPSPFVCLKIGACPSAH 480  
 Db 427 STKOEIIAALAEKGSFLPDPYQKODQFVAEYEPVLIEILVWMPSPFVCLKIGACPSAH 486  
 QY 481 KPLLGTEKCIWGPSYWCNTETAACNAVEHCKRHVNN 518  
 Db 487 KPLLGTEKCIWGPSYWCNTETAACNAVEHCKRHVNN 524  
 RESULT 3  
 AAY58716  
 ID AAY58716 standard; protein, 524 AA.  
 AC AAY58716;  
 DT 25-APR-2000 (first entry)  
 XX Human prosaposin.  
 DE  
 KW Prosaposin; saposin B; antiangiogenic; angiogenesis inhibitor;  
 KW antitumour; antiproliferative; antimigratory; Kaposi's sarcoma; tumour;  
 KW human; therapy.  
 XX Homo sapiens.  
 OS  
 XX Key Location/Qualifiers  
 FH 195..275  
 FT Protein /note= "mature saposin B"  
 FT Peptide 195..205  
 FT /note= "specifically claimed antiangiogenic peptide of  
 FT Claim 23"  
 FT 196..200  
 FT /note= "specifically claimed antiangiogenic peptide of  
 FT Claim 4"  
 XX WO200002902-A1.  
 XX 20-JAN-2000.  
 PD  
 XX 12-JUL-1999; 99WO-US015772.  
 PF  
 XX 13-JUL-1998; 98US-0092647P.  
 PR  
 XX (GILL/) GILL P S.  
 PA  
 XX GILL PS;  
 PI  
 XX WPI; 2000-171128/15.  
 DR  
 XX Saposin B derived peptides, useful as inhibitors of angiogenesis and  
 PT tumor growth.  
 PT  
 XX Disclosure; Page 18; 78pp; English.  
 XX  
 CC The present sequence is that of human prosaposin, a precursor of saposin  
 CC B. The invention is based on the discovery that saposin B, previously  
 CC known to be involved in the hydrolysis of sphingolipids, has potent  
 CC antiangiogenic and antitumour activity, and also has antiproliferative  
 CC and antimigratory activity against endothelial cells. This activity is  
 CC conserved in cryptic polypeptides as small as 5 amino acids (see AAY58684  
 CC -715), which can be synthetically prepared and used in vitro or in vivo  
 CC for the treatment of undesired angiogenesis and tumor growth, especially  
 CC Kaposi's sarcoma (claimed). The polypeptides can also be used in

CC conjunction with cytotoxic moieties to selectively kill certain cell  
 CC types, e.g. for treatment of cancer, angiodioma, neovascular glaucoma,  
 CC arteriovenous malformation, nonunion fracture, arthritis and other  
 CC connective tissue disorders, Osler-Weber syndrome, atherosclerotic  
 CC plaque, psoriasis, corneal graft neovascularization, pyogenic granuloma,  
 CC retrolental fibroplasia, diabetic retinopathy, scleroderma, haemangioma,  
 CC trachoma, vascular adhesions and hypertrophic scars  
 XX  
 SQ Sequence 524 AA;  
 Query Match 100.0%; Score 2759; DB 3; Length 524;  
 Best Local Similarity 100.0%; Pred. No. 7.2e-235;  
 Matches 518; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 LASLGAALAPVIGLKECTGSAVWCNVNTAADCAGVKGCLQTWANKPVPKSLPCIC 60  
 Db 7 LASLGAALAPVIGLKECTGSAVWCNVNTAADCAGVKGCLQTWANKPVPKSLPCIC 66  
 QY 61 KDVTTAAGDMLKDNATEEELVYLEKTCMDLPKPMMSASCKEIVDSYLPVLLDIKGEHS 120  
 Db 67 KDVTTAAGDMLKDNATEEELVYLEKTCMDLPKPMMSASCKEIVDSYLPVLLDIKGEHS 126  
 QY 121 RPEVCSALNLCESLQKHLAEINHQKLESNKIPELDMTEVAVAFPMANIPLLVPQDGP 180  
 Db 127 RPEVCSALNLCESLQKHLAEINHQKLESNKIPELDMTEVAVAFPMANIPLLVPQDGP 186  
 QY 181 SKPOKNDGVYCONCIQWNTDIOQAVRNSTFFVALVHVYEBECDRLGPMADICKNTIS 240  
 Db 187 SKPOKNDGVYCONCIQWNTDIOQAVRNSTFFVALVHVYEBECDRLGPMADICKNTIS 246  
 QY 241 QYSEIATOMMHMPKEICALVGFCDVEKEMPMOTLVPKASKNVLPALBELVEPIKHE 300  
 Db 247 QYSEIATOMMHMPKEICALVGFCDVEKEMPMOTLVPKASKNVLPALBELVEPIKHE 306  
 QY 301 VPAKSDVYCEVCEFLVKEVTKLIDNNKTEKEILDAFDMKCSKLPSLSSECOEVDVTGS 360  
 Db 307 VPAKSDVYCEVCEFLVKEVTKLIDNNKTEKEILDAFDMKCSKLPSLSSECOEVDVTGS 366  
 QY 361 SILSILAEVSPBELVCSMLHLCSGTRLPALTVHTVTPQPDGFCFCEVCKLVGLDRLNLEKN 420  
 Db 367 SILSILAEVSPBELVCSMLHLCSGTRLPALTVHTVTPQPDGFCFCEVCKLVGLDRLNLEKN 426  
 QY 421 STKOEIIAALAEKGSFLPDPYQKODQFVAEYEPVLIEILVWMPSPFVCLKIGACPSAH 480  
 Db 427 STKOEIIAALAEKGSFLPDPYQKODQFVAEYEPVLIEILVWMPSPFVCLKIGACPSAH 486  
 QY 481 KPLLGTEKCIWGPSYWCNTETAACNAVEHCKRHVNN 518  
 Db 487 KPLLGTEKCIWGPSYWCNTETAACNAVEHCKRHVNN 524  
 RESULT 4  
 ABUT9099  
 ID ABUT9099 standard; protein, 524 AA.  
 XX  
 AC ABUT9099;  
 DT 18-JUN-2003 (first entry)  
 XX  
 DE Lip-TAA binding protein, Prosaposin.  
 XX  
 KW Superantigen; SAg; staphylococcal enterotoxin; tumour; cancer; apoptosis;  
 KW gene therapy; mammalian cell receptor; cytostatic;  
 KW tumour associated lipid; energy; T cell; antigen presenting cell; APC;  
 KW tumouricidal immunocyte; antitumour.  
 XX  
 OS Unidentified.  
 XX US2002177551-A1.  
 XX 28-NOV-2002.  
 PD 30-MAY-2001; 2001US-00870759.  
 PF

31-MAY-2000; 2000US-0208128P.  
(TERM/) TERMAN D S.  
Terman DS;  
WPI; 2003-361759/34.  
A mammalian cell receptor, useful in the treatment of cancer by binding to tumor associated lipids where the binding induces anergy or apoptosis in T cells and antigen presenting cells.  
Disclosure; Page; 167pp; English.  
The invention relates to a mammalian cell receptor, useful in the treatment of cancer, which binds to tumour associated lipids and induces anergy or apoptosis in the T cells and antigen presenting cells (APCs). Also included are a mammalian cell useful in the treatment of cancer where the receptor which binds tumour associated lipids and induces cellular inactivation or death is deleted or functionally deactivated, producing (M1) a tumouricidal immunocyte population in vivo in a mammal by allowing tumour associated lipids to contact immunocytes in which receptors for immunosuppressive fatty acids, ceramides, glycolipids, sphingolipids, glycosphingolipid, phosphoglycolipids, gangliosides, sialylated glycans, lipopeptides and phospholipolipids are inactivated or deleted, a construct useful in the treatment of cancer comprising a superantigen (SAg) nucleotide inserted into a virus, a mammalian T cell useful in the treatment of cancer (where an adaptor protein which inhibits T cell activation by tumour associated antigens is deleted or functionally deactivated), a composition useful in the treatment of cancer (comprising a lipid raft conjugated to a superantigen), producing (M2) a tumouricidal immunocyte population ex vivo in a mammal (by allowing tumour associated lipid to contact immunocytes, in which receptors for the lipids are inactivated or deleted to produce a tumouricidal immunocyte population, and administering the tumouricidally activated immunocytes to the host), producing (M3) a tumouricidal APC population ex vivo in a mammal (by allowing a tumour associated lipid to contact APCs, in which receptors for the tumour associated lipids are inactivated or deleted to produce a tumouricidally activated population, and administering APCs to the host), producing a tumouricidal T cell population ex vivo in a mammal) by allowing a tumour associated lipid to contact T cells, in which adaptor proteins, which inhibit T cell activation by tumour associated antigens, are deleted or functionally deactivated to produce a tumouricidal population of T cells, and administering the tumouricidally activated T cells to the host, or allowing a superantigen-lipid raft to contact T cells ex vivo, and administering the tumouricidally activated T cells to the host, and (M5) cancer in a mammal (by administering a lipid binding molecule which binds immunosuppressive tumour associated lipids in vivo), producing (M6) a tumouricidal T cell population in vivo in a mammal (by allowing a tumour associated antigen to contact immunocytes in which adaptor proteins which inhibit T cell activation by tumour associated antigens are deleted or functionally deactivated) and producing (M7) a tumouricidal T cell population ex vivo in a mammal comprising allowing a superantigen-lipid raft conjugate to contact immunocytes in vivo. The receptors, methods and compositions are useful for treating cancers and tumours. Bacterial superantigens are co-administered or administered as fusion constructs with anti-tumour proteins or motifs. The present sequence represents a tumour antigen or a motif identifying a tumour antigen, which can be functionally deactivated in the method of the invention. Note: The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format from the US patent office website at "seqdata.uspto.gov/sequence.html?DocID=20020177551"

Query Match	100.0%;	Score 2759;	DB 6;	Length 524;
Best Local Similarity	100.0%;	Pred. No. 7,2e-235;		
Matches 518;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;

[illegible]

PT New polypeptides (e.g. kinases, phosphatases, proteases, transporters, cytoskeletal proteins, receptors or transcription factors), useful for treating cancer, e.g. colon cancer, gastric cancer, sarcoma, lymphoma or leukemia.

Example 2; SEQ ID NO 1866; 134pp; English.

The invention describes a purified polypeptide which comprises a fragment of a kinase, phosphatase, protease, protease inhibitor, transporter, cytoskeletal protein, receptor or transcription factor. The polypeptide is useful as an immunogenic composition for eliciting in a mammal an immunogenic response directed against any of the purified polypeptide, the purified polypeptide, or the antibody that binds to this polypeptide, is useful for treating cancer. The polypeptide is also useful for identifying compounds that binds to a naturally processed class I or class II MHC-binding polypeptide. The polypeptides and polynucleotides are particularly useful for treating or preventing melanoma, colon cancer, gastric cancer, adenocarcinoma, sarcoma, melanoma, lymphoma or leukemia. These are also useful for screening agents for treating the above mentioned diseases. This sequence represents an expressed protein tag (EPT) isolated from human tissue for translational profiling. Note: This sequence does not appear in the printed specification but was obtained in electronic format directly from WIPO at [http://wipo.int/pub/publ/pat\\_sequences](http://wipo.int/pub/publ/pat_sequences)

SQ Sequence 524 AA;

Query Match	100.0%	Score 2759	DB 6	Length 524
Best Local Similarity	100.0%	Pred. NO. 7.2e-235		
Matches 518; Conservative	0	Mismatches 0	Indels 0	Gaps 0

QY	1	LASLLGALAGAVLGLIKECTRGSAVWCQVVKTSJSDCAVGHCIQTWNNKFTVSLPDCIC	60
Db	7	LASLLGALAGAVLGLIKECTRGSAVWCQVVKTSJSDCAVGHCIQTWNNKFTVSLPDCIC	66
QY	61	KDVYTAAGM.LKDNATEEBILVYLEKTCIMLPPNNNASCKEIVDSYLPYITDI IKGMS	120
Db	67	KDVYTAAGM.LKDNATEEBILVYLEKTCIMLPPNNNASCKEIVDSYLPYITDI IKGMS	126
QY	121	RPGEVCSALNLCESLQKHLAEINHQKLSNKKPELDMTEVVA.PFMANIP.LLLLYPODGR	180
Db	127	RPGEVCSALNLCESLQKHLAEINHQKLSNKKPELDMTEVVA.PFMANIP.LLLLYPODGR	186
QY	181	SKPQPKDNGDV.CODCICQWVTDIQTAVRINSTFVQALVEHYKECDRLGPGMADICKNYIS	240
Db	187	SKPQPKDNGDV.CODCICQWVTDIQTAVRINSTFVQALVEHYKECDRLGPGMADICKNYIS	246
QY	241	OYSELAIQMMMHMOPKEICALVGFCDQEVEMPMQTLVPAKVASKNYIPALEVEPIIKKE	300
Db	247	OYSELAIQMMMHMOPKEICALVGFCDQEVEMPMQTLVPAKVASKNYIPALEVEPIIKKE	306
QY	301	VPAKSDVYCEVCEFLVKEVTKLIDNNKTEKEILDAFDKMSCKLPSLSBECQEVVDYGS	360
Db	307	VPAKSDVYCEVCEFLVKEVTKLIDNNKTEKEILDAFDKMSCKLPSLSBECQEVVDYGS	366
QY	361	STLSLILBEVSPBLYCSM.LHCSGTRLPALTVAVTQPKQGFCEVCCKLVGYIDRNLKKN	420
Db	367	STLSLILBEVSPBLYCSM.LHCSGTRLPALTVAVTQPKQGFCEVCCKLVGYIDRNLKKN	426
QY	421	STKQELIAL.EKGCSFLPDYQKQCOQFAAEYEPVLIILVEVMDPSFVCLKIGACPSAH	480
Db	427	STKQELIAL.EKGCSFLPDYQKQCOQFAAEYEPVLIILVEVMDPSFVCLKIGACPSAH	486
QY	481	KPLLTGTEKCIWGPSTWCQNTETAAQCNAAVEHCKRHHWN	518
Db	487	KPLLTGTEKCIWGPSTWCQNTETAAQCNAAVEHCKRHHWN	524

RESULT 6  
ABU05207  
ID ABU05207 standard; protein; 524 AA  
XX  
AC ABU05207;

XX	29-JAN-2003 (first entry)
DT	
XX	
DE	Human expressed protein t

KM Translational profiling; expressed protein tag; EPT; kinase; phosphatase  
 KM protease; protease inhibitor; transporter; cytoskeletal protein;  
 KM receptor; transcription factor; cancer; MHC;  
 KM major histocompatibility complex; myeloma; colon cancer; gastric cancer;  
 KM adenocarcinoma; sarcoma; melanoma; lymphoma; leukaemia.

OS Homo sapiens.

PN WO200278524-A2

PD 10-OCT-2002.

PF 28-MAR-2002; 2002WO-US009671.  
XY

PR 28-MAR-2001; 2001US-02/9495P.  
PR 31-MAY-2001; 2001TIS-0292544P

01-OCT-2001: 2001HS-0326370P  
PK 08-AUG-2001; 2001HS-0310801P  
PB

04-DEC-2001; 200108-0338/80F  
PR 20-FEB-2002: 200205-0358985P

(ZYCO-) ZYCOS INC,  
PA

PI Chicz RM, Tomlinson AJ, Urth

DR WPI; 2003-040607/03.

PT New polypeptides

PT treating cancer, e.g. colon cancer,

1  
2  
3  
4  
5  
6  
7  
8  
9  
10  
11  
12  
13  
14  
15  
16  
17  
18  
19  
20  
21  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60  
61  
62  
63  
64  
65  
66  
67  
68  
69  
70  
71  
72  
73  
74  
75  
76  
77  
78  
79  
80  
81  
82  
83  
84  
85  
86  
87  
88  
89  
90  
91  
92  
93  
94  
95  
96  
97  
98  
99  
100  
101  
102  
103  
104  
105  
106  
107  
108  
109  
110  
111  
112  
113  
114  
115  
116  
117  
118  
119  
120  
121  
122  
123  
124  
125  
126  
127  
128  
129  
130  
131  
132  
133  
134  
135  
136  
137  
138  
139  
140  
141  
142  
143  
144  
145  
146  
147  
148  
149  
150  
151  
152  
153  
154  
155  
156  
157  
158  
159  
160  
161  
162  
163  
164  
165  
166  
167  
168  
169  
170  
171  
172  
173  
174  
175  
176  
177  
178  
179  
180  
181  
182  
183  
184  
185  
186  
187  
188  
189  
190  
191  
192  
193  
194  
195  
196  
197  
198  
199  
200  
201  
202  
203  
204  
205  
206  
207  
208  
209  
210  
211  
212  
213  
214  
215  
216  
217  
218  
219  
220  
221  
222  
223  
224  
225  
226  
227  
228  
229  
230  
231  
232  
233  
234  
235  
236  
237  
238  
239  
240  
241  
242  
243  
244  
245  
246  
247  
248  
249  
250  
251  
252  
253  
254  
255  
256  
257  
258  
259  
260  
261  
262  
263  
264  
265  
266  
267  
268  
269  
270  
271  
272  
273  
274  
275  
276  
277  
278  
279  
280  
281  
282  
283  
284  
285  
286  
287  
288  
289  
290  
291  
292  
293  
294  
295  
296  
297  
298  
299  
300  
301  
302  
303  
304  
305  
306  
307  
308  
309  
310  
311  
312  
313  
314  
315  
316  
317  
318  
319  
320  
321  
322  
323  
324  
325  
326  
327  
328  
329  
330  
331  
332  
333  
334  
335  
336  
337  
338  
339  
340  
341  
342  
343  
344  
345  
346  
347  
348  
349  
350  
351  
352  
353  
354  
355  
356  
357  
358  
359  
360  
361  
362  
363  
364  
365  
366  
367  
368  
369  
370  
371  
372  
373  
374  
375  
376  
377  
378  
379  
380  
381  
382  
383  
384  
385  
386  
387  
388  
389  
390  
391  
392  
393  
394  
395  
396  
397  
398  
399  
400  
401  
402  
403  
404  
405  
406  
407  
408  
409  
410  
411  
412  
413  
414  
415  
416  
417  
418  
419  
420  
421  
422  
423  
424  
425  
426  
427  
428  
429  
430  
431  
432  
433  
434  
435  
436  
437  
438  
439  
440  
441  
442  
443  
444  
445  
446  
447  
448  
449  
450  
451  
452  
453  
454  
455  
456  
457  
458  
459  
460  
461  
462  
463  
464  
465  
466  
467  
468  
469  
470  
471  
472  
473  
474  
475  
476  
477  
478  
479  
480  
481  
482  
483  
484  
485  
486  
487  
488  
489  
490  
491  
492  
493  
494  
495  
496  
497  
498  
499  
500  
501  
502  
503  
504  
505  
506  
507  
508  
509  
510  
511  
512  
513  
514  
515  
516  
517  
518  
519  
520  
521  
522  
523  
524  
525  
526  
527  
528  
529  
530  
531  
532  
533  
534  
535  
536  
537  
538  
539  
540  
541  
542  
543  
544  
545  
546  
547  
548  
549  
550  
551  
552  
553  
554  
555  
556  
557  
558  
559  
560  
561  
562  
563  
564  
565  
566  
567  
568  
569  
570  
571  
572  
573  
574  
575  
576  
577  
578  
579  
580  
581  
582  
583  
584  
585  
586  
587  
588  
589  
590  
591  
592  
593  
594  
595  
596  
597  
598  
599  
600  
601  
602  
603  
604  
605  
606  
607  
608  
609  
610  
611  
612  
613  
614  
615  
616  
617  
618  
619  
620  
621  
622  
623  
624  
625  
626  
627  
628  
629  
630  
631  
632  
633  
634  
635  
636  
637  
638  
639  
640  
641  
642  
643  
644  
645  
646  
647  
648  
649  
650  
651  
652  
653  
654  
655  
656  
657  
658  
659  
660  
661  
662  
663  
664  
665  
666  
667  
668  
669  
670  
671  
672  
673  
674  
675  
676  
677  
678  
679  
680  
681  
682  
683  
684  
685  
686  
687  
688  
689  
690  
691  
692  
693  
694  
695  
696  
697  
698  
699  
700  
701  
702  
703  
704  
705  
706  
707  
708  
709  
710  
711  
712  
713  
714  
715  
716  
717  
718  
719  
720  
721  
722  
723  
724  
725  
726  
727  
728  
729  
730  
731  
732  
733  
734  
735  
736  
737  
738  
739  
740  
741  
742  
743  
744  
745  
746  
747  
748  
749  
750  
751  
752  
753  
754  
755  
756  
757  
758  
759  
760  
761  
762  
763  
764  
765  
766  
767  
768  
769  
770  
771  
772  
773  
774  
775  
776  
777  
778  
779  
780  
781  
782  
783  
784  
785  
786  
787  
788  
789  
790  
791  
792  
793  
794  
795  
796  
797  
798  
799  
800  
801  
802  
803  
804  
805  
806  
807  
808  
809  
810  
811  
812  
813  
814  
815  
816  
817  
818  
819  
820  
821  
822  
823  
824  
825  
826  
827  
828  
829  
830  
831  
832  
833  
834  
835  
836  
837  
838  
839  
840  
84

1  
 2  
 3  
 4  
 5  
 6  
 7  
 8  
 9  
 10  
 11  
 12  
 13  
 14  
 15  
 16  
 17  
 18  
 19  
 20  
 21  
 22  
 23  
 24  
 25  
 26  
 27  
 28  
 29  
 30  
 31  
 32  
 33  
 34  
 35  
 36  
 37  
 38  
 39  
 40  
 41  
 42  
 43  
 44  
 45  
 46  
 47  
 48  
 49  
 50  
 51  
 52  
 53  
 54  
 55  
 56  
 57  
 58  
 59  
 60  
 61  
 62  
 63  
 64  
 65  
 66  
 67  
 68  
 69  
 70  
 71  
 72  
 73  
 74  
 75  
 76  
 77  
 78  
 79  
 80  
 81  
 82  
 83  
 84  
 85  
 86  
 87  
 88  
 89  
 90  
 91  
 92  
 93  
 94  
 95  
 96  
 97  
 98  
 99  
 100  
 101  
 102  
 103  
 104  
 105  
 106  
 107  
 108  
 109  
 110  
 111  
 112  
 113  
 114  
 115  
 116  
 117  
 118  
 119  
 120  
 121  
 122  
 123  
 124  
 125  
 126  
 127  
 128  
 129  
 130  
 131  
 132  
 133  
 134  
 135  
 136  
 137  
 138  
 139  
 140  
 141  
 142  
 143  
 144  
 145  
 146  
 147  
 148  
 149  
 150  
 151  
 152  
 153  
 154  
 155  
 156  
 157  
 158  
 159  
 160  
 161  
 162  
 163  
 164  
 165  
 166  
 167  
 168  
 169  
 170  
 171  
 172  
 173  
 174  
 175  
 176  
 177  
 178  
 179  
 180  
 181  
 182  
 183  
 184  
 185  
 186  
 187  
 188  
 189  
 190  
 191  
 192  
 193  
 194  
 195  
 196  
 197  
 198  
 199  
 200  
 201  
 202  
 203  
 204  
 205  
 206  
 207  
 208  
 209  
 210  
 211  
 212  
 213  
 214  
 215  
 216  
 217  
 218  
 219  
 220  
 221  
 222  
 223  
 224  
 225  
 226  
 227  
 228  
 229  
 230  
 231  
 232  
 233  
 234  
 235  
 236  
 237  
 238  
 239  
 240  
 241  
 242  
 243  
 244  
 245  
 246  
 247  
 248  
 249  
 250  
 251  
 252  
 253  
 254  
 255  
 256  
 257  
 258  
 259  
 260  
 261  
 262  
 263  
 264  
 265  
 266  
 267  
 268  
 269  
 270  
 271  
 272  
 273  
 274  
 275  
 276  
 277  
 278  
 279  
 280  
 281  
 282  
 283  
 284  
 285  
 286  
 287  
 288  
 289  
 290  
 291  
 292  
 293  
 294  
 295  
 296  
 297  
 298  
 299  
 300  
 301  
 302  
 303  
 304  
 305  
 306  
 307  
 308  
 309  
 310  
 311  
 312  
 313  
 314  
 315  
 316  
 317  
 318  
 319  
 320  
 321  
 322  
 323  
 324  
 325  
 326  
 327  
 328  
 329  
 330  
 331  
 332  
 333  
 334  
 335  
 336  
 337  
 338  
 339  
 340  
 341  
 342  
 343  
 344  
 345  
 346  
 347  
 348  
 349  
 350  
 351  
 352  
 353  
 354  
 355  
 356  
 357  
 358  
 359  
 360  
 361  
 362  
 363  
 364  
 365  
 366  
 367  
 368  
 369  
 370  
 371  
 372  
 373  
 374  
 375  
 376  
 377  
 378  
 379  
 380  
 381  
 382  
 383  
 384  
 385  
 386  
 387  
 388  
 389  
 390  
 391  
 392  
 393  
 394  
 395  
 396  
 397  
 398  
 399  
 400  
 401  
 402  
 403  
 404  
 405  
 406  
 407  
 408  
 409  
 410  
 411  
 412  
 413  
 414  
 415  
 416  
 417  
 418  
 419  
 420  
 421  
 422  
 423  
 424  
 425  
 426  
 427  
 428  
 429  
 430  
 431  
 432  
 433  
 434  
 435  
 436  
 437  
 438  
 439  
 440  
 441  
 442  
 443  
 444  
 445  
 446  
 447  
 448  
 449  
 450  
 451  
 452  
 453  
 454  
 455  
 456  
 457  
 458  
 459  
 460  
 461  
 462  
 463  
 464  
 465  
 466  
 467  
 468  
 469  
 470  
 471  
 472  
 473  
 474  
 475  
 476  
 477  
 478  
 479  
 480  
 481  
 482  
 483  
 484  
 485  
 486  
 487  
 488  
 489  
 490  
 491  
 492  
 493  
 494  
 495  
 496  
 497  
 498  
 499  
 500  
 501  
 502  
 503  
 504  
 505  
 506  
 507  
 508  
 509  
 510  
 511  
 512  
 513  
 514  
 515  
 516  
 517  
 518  
 519  
 520  
 521  
 522  
 523  
 524  
 525

The invention describes a purified polypeptide which comprises a fragment of a kinase, phosphatase, protease, protease inhibitor, transporter, cytoskeletal protein, receptor or transcription factor. The polypeptide is useful as an immunogenic composition for eliciting in a mammal an immunogenic response directed against any of the purified polypeptide. The purified polypeptide, or the antibody that binds to this polypeptide, is useful for treating cancer. The polypeptide is also useful for identifying compounds that binds to a naturally processed class I or class II MHC-binding polypeptide. The polypeptides and polynucleotides are particularly useful for treating or preventing myeloma, colon cancer, gastric cancer, adenocarcinoma, sarcoma, melanoma, lymphoma or leukemia. These are also useful for screening agents for treating the above mentioned diseases. This sequence represents an expressed protein tag (EPT) isolated from human tissue for translational profiling. Note: This sequence does not appear in the printed specification but was obtained in electronic format directly from WIPO at [http://wipo.int/pub/published\\_pct\\_sequences](http://wipo.int/pub/published_pct_sequences)

**SQ Sequence 524 AA;**

Query Match	100.0%;	Score 2759;	DB 6;	Length 534;
Best Local Similarity	100.0%;	Pred. No. 7	2e-235;	
Matches 518; Conservative	0;	Mismatches	0;	Indels 0
				Gaps 0

QY 1 LASSLGAALAGVLTGKRECTRGSSAWCQWVKIASDCAVHCHCIQYWNKRPVSLPDDIC 60

Db 7 LASSLGAALAGVLTGKRECTRGSSAWCQWVKIASDCAVHCHCIQYWNKRPVSLPDDIC 66

QY 61 KDVTAAQDMLKDNATEEBEILVLEKTCMLPRPNNSASCKEIVDSYLPVITDI IKGEMS 120

Db 67 KDVTAAQDMLKDNATEEBEILVLEKTCMLPRPNNSASCKEIVDSYLPVITDI IKGEMS 126

QY 121 RGEVCSALNTCESIQKHIAELNHQKQLESNNKIPSLDMEVVAFEMANTPLLLYPODGP 180

Db 127 RPEVCSALNLCESIQKHLAEINHQLESNKIPELDMTEVAVAFEMANIPILLYPQDPR 186  
 QY 181 SKPQKNDGVDCODCIQWVTDIQTAVRTNSTFVQALVEHYVEECRLPGMADICKNYIS 240  
 Db 187 SKPQKNDGVDCODCIQWVTDIQTAVRTNSTFVQALVEHYVEECRLPGMADICKNYIS 246  
 QY 241 QYSEIATQMMHMOPEKICALVGCDEVKEMPMQTLVPAKVASKNVLPALVEPIKKEH 300  
 Db 247 QYSEIATQMMHMOPEKICALVGCDEVKEMPMQTLVPAKVASKNVLPALVEPIKKEH 306  
 QY 301 VPAKSDVYCEVCEFLVKEVTKLIDNNKTEKEIIDAFDMCKSLPKLSIEECQEVVTYGS 360  
 Db 307 VPAKSDVYCEVCEFLVKEVTKLIDNNKTEKEIIDAFDMCKSLPKLSIEECQEVVTYGS 366  
 QY 361 SILSILIEEVSPELVCSMLHLCSGTRLPALTVAHTQPKDGFCEVCKLVGYIDRLNLEKN 420  
 Db 367 SILSILIEEVSPELVCSMLHLCSGTRLPALTVAHTQPKDGFCEVCKLVGYIDRLNLEKN 426  
 QY 421 STKQEIILALEKGCSEFLPDYQKQCDQFVAEYEPVLIEILVEVMDPSFVCLIKIGACPSAH 480  
 Db 427 STKQEIILALEKGCSEFLPDYQKQCDQFVAEYEPVLIEILVEVMDPSFVCLIKIGACPSAH 486  
 QY 481 KPLGTGTEKCIWGPSSYWCQNTETTAACNAVEHCKRHVWN 518  
 Db 487 KPLGTGTEKCIWGPSSYWCQNTETTAACNAVEHCKRHVWN 524

## RESULT 7

ABU05203  
 ID ABU05203 standard; protein; 524 AA.

AC ABU05203;

DT 29-JAN-2003 (first entry)

DE Human expressed protein tag (EPT) #1869.

KW Translational profiling; expressed protein tag; EPT; kinase; phosphatase;

KW protease; protease inhibitor; transporter; cytoskeletal protein;

KW receptor; transcription factor; cancer; MHC;

KW major histocompatibility complex; myeloma; colon cancer; gastric cancer;

KW adenocarcinoma; sarcoma; melanoma; lymphoma; leukemia.

OS Homo sapiens.

PN MO200278524-A2.

PD 10-OCT-2002.

PF 28-MAR-2002; 2002MO-US009671.

PR 28-MAR-2001; 2001US-0279495P.

PR 21-MAY-2001; 2001US-0292544P.

PR 08-AUG-2001; 2001US-0310801P.

PR 01-OCT-2001; 2001US-0326370P.

PR 04-DEC-2001; 2001US-0336780P.

PR 20-FEB-2002; 2002US-0358985P.

PA (ZYCO-) ZYCOS INC.

PI Chiciz RM, Tomlinson AJ, Urban RG;

DR WPI; 2003-040607/03.

PT New polypeptides (e.g. kinases, phosphatases, proteases, transporters,

PT cytoskeletal protein, receptors or transcription factors), useful for

PT treating cancer, e.g. colon cancer, gastric cancer, sarcoma, lymphoma or

PT leukemia.

CC Example 2; SEQ ID NO 1869; 134pp; English.

CC The invention describes a purified polypeptide, which comprises a

CC fragment of a kinase, phosphatase, protease, protease inhibitor,

CC transporter, cytoskeletal protein, receptor or transcription factor. The

CC polypeptide is useful as an immunogenic composition for eliciting in a

CC mammal an immunogenic response directed against any of the purified

CC polypeptide. The purified polypeptide, or the antibody that binds to this

CC polypeptide, is useful for treating cancer. The polypeptide is also

CC useful for identifying compounds that bind to a naturally processed

CC class I or class II MHC-binding polypeptide. The polypeptides and

CC polynucleotides are particularly useful for treating or preventing

CC myeloma, colon cancer, gastric cancer, adenocarcinoma, sarcoma, melanoma,

CC lymphoma or leukemia. These are also useful for screening agents for

CC treating the above mentioned diseases. This sequence represents an

CC expressed protein tag (EPT) isolated from human tissue for translational

CC profiling. Note: This sequence does not appear in the printed

CC specification but was obtained in electronic format directly from WIP0 at

CC ftp.wipo.int/pub/published\_pct\_sequences

CC Sequence 524 AA:

CC Query Match 100.0%; Score 2759; DB 6; Length 524;

CC Best Local Similarity 100.0%; Pred. No. 7.2e-235;

CC Matches 518; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LASTLGAALGAPVLGLKECTRGSAVWCQVKTASDGAAYHCLQTVWKNKPTVKSIPCDIC 60

Db 7 LASTLGAALGAPVLGLKECTRGSAVWCQVKTASDGAAYHCLQTVWKNKPTVKSIPCDIC 66

QY 61 KDVTAAAGMDLKNATEEELVLEKTCMDLPRPNNSASCKEIVDSYPIVLLDIIKGEWS 120

Db 67 KDVTAAAGMDLKNATEEELVLEKTCMDLPRPNNSASCKEIVDSYPIVLLDIIKGEWS 126

QY 121 RPEVCSALNLCESIQKHLAEINHQLESNKIPELDMTEVAVAFEMANIPILLYPQDPR 180

Db 127 RPEVCSALNLCESIQKHLAEINHQLESNKIPELDMTEVAVAFEMANIPILLYPQDPR 186

QY 181 SKPQKNDGVDCODCIQWVTDIQTAVRTNSTFVQALVEHYVEECRLPGMADICKNYIS 240

Db 187 SKPQKNDGVDCODCIQWVTDIQTAVRTNSTFVQALVEHYVEECRLPGMADICKNYIS 246

QY 241 QYSEIATQMMHMOPEKICALVGCDEVKEMPMQTLVPAKVASKNVLPALVEPIKKEH 300

Db 247 QYSEIATQMMHMOPEKICALVGCDEVKEMPMQTLVPAKVASKNVLPALVEPIKKEH 306

QY 301 VPAKSDVYCEVCEFLVKEVTKLIDNNKTEKEIIDAFDMCKSLPKLSIEECQEVVTYGS 360

Db 307 VPAKSDVYCEVCEFLVKEVTKLIDNNKTEKEIIDAFDMCKSLPKLSIEECQEVVTYGS 366

QY 361 SILSILIEEVSPELVCSMLHLCSGTRLPALTVAHTQPKDGFCEVCKLVGYIDRLNLEKN 420

Db 367 SILSILIEEVSPELVCSMLHLCSGTRLPALTVAHTQPKDGFCEVCKLVGYIDRLNLEKN 426

QY 421 STKQEIILALEKGCSEFLPDYQKQCDQFVAEYEPVLIEILVEVMDPSFVCLIKIGACPSAH 480

Db 427 STKQEIILALEKGCSEFLPDYQKQCDQFVAEYEPVLIEILVEVMDPSFVCLIKIGACPSAH 486

QY 481 KPLGTGTEKCIWGPSSYWCQNTETTAACNAVEHCKRHVWN 518

Db 487 KPLGTGTEKCIWGPSSYWCQNTETTAACNAVEHCKRHVWN 524

RESULT 8

ABU07340

ID ABU07340 standard; protein; 524 AA.

AC ABU07340;

DT 29-JAN-2003 (first entry)

DE Human expressed protein tag (EPT) #2041.

KW Translational profiling; expressed protein tag; EPT; kinase; phosphatase;

KW protease; protease inhibitor; transporter; cytoskeletal protein;

KW receptor; transcription factor; cancer; MHC;

KW major histocompatibility complex; myeloma; colon cancer; gastric cancer;



KW adenocarcinoma; sarcoma; melanoma; lymphoma; leukemia.  
 XX Homo sapiens.  
 OS  
 PN MO200278524-A2.  
 XX  
 XX 10-OCT-2002.  
 PD  
 XX 28-MAR-2002; 2002MO-US009671.  
 PR  
 XX 28-MAR-2001; 2001US-0279495P.  
 PR 21-MAY-2001; 2001US-0292544P.  
 PR 08-AUG-2001; 2001US-0310801P.  
 PR 01-OCT-2001; 2001US-0326370P.  
 PR 04-DEC-2001; 2001US-0336780P.  
 PR 20-FEB-2002; 2002US-0358985P.  
 XX  
 XX (ZYCO-) ZYCOS INC.  
 PA  
 PI Chiciz RM, Tomlinson AJ, Urban RG;  
 XX  
 XX WPI; 2003-040607/03.  
 DR  
 XX  
 XX New polypeptides (e.g. kinases, phosphatases, proteases, transporters,  
 PT cytoskeletal proteins, receptors or transcription factors), useful for  
 PT treating cancer, e.g. colon cancer, gastric cancer, sarcoma, lymphoma or  
 PT leukemia.  
 XX  
 XX Example 2; SEQ ID NO 2041; 134pp; English.  
 PS  
 XX The invention describes a purified polypeptide, which comprises a  
 CC fragment of a kinase, phosphatase, protease, or transcription inhibitor,  
 CC transporter, cytoskeletal protein, receptor or transcription factor. The  
 CC polypeptide is useful as an immunogenic composition for eliciting in a  
 CC mammal an immunogenic response directed against any of the purified  
 CC polypeptide. The purified polypeptide, or the antibody that binds to this  
 CC polypeptide, is useful for treating cancer. The polypeptide is also  
 CC useful for identifying compounds that binds to a naturally processed  
 CC class I or class II MHC-binding polypeptide. The polypeptides and  
 CC polynucleotides are particularly useful for treating or preventing  
 CC myeloma, colon cancer, gastric cancer, adenocarcinoma, sarcoma, melanoma,  
 CC lymphoma or leukemia. These are also useful for screening agents for  
 CC treating the above mentioned diseases. This sequence represents an  
 CC expressed protein tag (EPT) isolated from human tissue for translational  
 CC profiling. Note: This sequence does not appear in the printed  
 CC specification but was obtained in electronic format directly from WIPO at  
 CC ftp.wipo.int/pub/published\_pct\_sequences  
 CC  
 XX  
 XX Sequence 524 AA:  
 SQ  
 Query Match 100.0%; Score 2759; DB 6; Length 524;  
 Best Local Similarity 100.0%; Pred. No. 7,2e-235;  
 Matches 518; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 301 VPAKSDVYCEVCEFLAVKVTLLIDNNKTEKILLDAFDKMGSKLPKSLSESGQEVVDYGS 360  
 DB 307 VPAKSDVYCEVCEFLAVKVTLLIDNNKTEKILLDAFDKMGSKLPKSLSESGQEVVDYGS 366  
 QY 361 SILSILLEEVSPELVCSMLHLCSGSTRPALVHTVTPKDGSCFCEYCKKLAVGLDRLNLEKN 420  
 DB 367 SILSILLEEVSPELVCSMLHLCSGSTRPALVHTVTPKDGSCFCEYCKKLAVGLDRLNLEKN 426  
 QY 421 STKOEIILAALKGCSFLPDPYQKODQFVAEYEVLLIILVWDPSPVCLKIGACPSAH 480  
 DB 427 STKOEIILAALKGCSFLPDPYQKODQFVAEYEVLLIILVWDPSPVCLKIGACPSAH 486  
 QY 481 KPLIGTEKCIWGPSTWCONTEETAACNAVEHCKRHWN 518  
 DB 487 KPLIGTEKCIWGPSTWCONTEETAACNAVEHCKRHWN 524

RESULT 9  
 ABU05216  
 ID ABU05216 standard; protein; 524 AA.  
 XX  
 XX AC ABU05216;  
 XX  
 XX 29-JAN-2003 (first entry)  
 DT  
 XX  
 XX Human expressed protein tag (EPT) #1882.  
 DB  
 XX  
 XX Translational profiling; expressed protein tag; EPT; kinase, phosphatase;  
 KW protease; protease inhibitor; transporter; cytoskeletal protein;  
 KW receptor; transcription factor; cancer; MHC;  
 KW major histocompatibility complex; myeloma; colon cancer; gastric cancer;  
 KW adenocarcinoma; sarcoma; melanoma; lymphoma; leukemia.  
 XX  
 OS Homo sapiens.  
 XX  
 PN MO200278524-A2.  
 PD  
 XX 10-OCT-2002.  
 PD  
 XX 28-MAR-2002; 2002MO-US009671.  
 PF  
 XX 28-MAR-2001; 2001US-0279495P.  
 PR 21-MAY-2001; 2001US-0292544P.  
 PR 08-AUG-2001; 2001US-0310801P.  
 PR 01-OCT-2001; 2001US-0326370P.  
 PR 04-DEC-2001; 2001US-0336780P.  
 PR 20-FEB-2002; 2002US-0358985P.  
 XX  
 XX (ZYCO-) ZYCOS INC.  
 PA  
 PI Chiciz RM, Tomlinson AJ, Urban RG;  
 XX  
 XX WPI; 2003-040607/03.  
 DR  
 XX  
 XX New polypeptides (e.g. kinases, phosphatases, proteases, transporters,  
 PT cytoskeletal proteins, receptors or transcription factors), useful for  
 PT treating cancer, e.g. colon cancer, gastric cancer, sarcoma, lymphoma or  
 PT leukemia.  
 XX  
 XX Example 2; SEQ ID NO 1882; 134pp; English.  
 PS  
 XX The invention describes a purified polypeptide, which comprises a  
 CC fragment of a kinase, phosphatase, protease, or transcription inhibitor,  
 CC transporter, cytoskeletal protein, receptor or transcription factor. The  
 CC polypeptide is useful as an immunogenic composition for eliciting in a  
 CC mammal an immunogenic response directed against any of the purified  
 CC polypeptide. The purified polypeptide, or the antibody that binds to this  
 CC polypeptide, is useful for treating cancer. The polypeptide is also  
 CC useful for identifying compounds that binds to a naturally processed  
 CC class I or class II MHC-binding polypeptide. The polypeptides and  
 CC polynucleotides are particularly useful for treating or preventing  
 CC myeloma, colon cancer, gastric cancer, adenocarcinoma, sarcoma, melanoma,

CC lymphoma or leukaemia. These are also useful for screening agents for  
CC treating the above mentioned diseases. This sequence represents an  
CC expressed protein tag (EPT) isolated from human tissue for translational  
CC profiling. Note: This sequence does not appear in the printed  
CC specification but was obtained in electronic format directly from WIPD at  
CC ftp.wipo.int/pub/published\_pct\_sequences  
XX  
SQ Sequence 524 Aa;

Query Match	100.0%	Score 2759;	DB 6;	Length 524;
Best Local Similarity	100.0%	Pred. No. 7.2e-235;		
Matches 518; Conservative	0;	Mismatches	0;	Gap

Qy	7	LASTLGGALAPRYGLKECTRGSVWCONKTSDDGAVKGLQTYWNPKYTSJPCDTC	60
Db	7	LASLIGALAPRYGLKECTRGSVWQONVKTASDGAHVGLQTYWNPKYTSJPCDTC	66
Qy	61	KDVTYAGDMLKDNATEEELLVYLEKTCDWLPKPNNSASCKEIVDSYLPVLIIDIIKGEMS	120
Db	67	KDVTYAGDMLKDNATEEELLVYLEKTCDWLPKPNNSASCKEIVDSYLPVLIIDIIKGEMS	128
Qy	121	RPEVCSALNLCESLQHLAEIHNQOLBENKIPBLDMTEVYAPPMANIPLILYPDQGR	180
Db	127	RPEVCSALNLCESLQHLAEIHNQOLBENKIPBLDMTEVYAPPMANIPLILYPDQGR	186
Qy	181	SKPQPKONGVQDCDQIMQVTDIQTAVRINSTFOALVHEVHECECRGLRPGMADI	240
Db	187	SKPQPKONGVQDCDQIMQVTDIQTAVRINSTFOALVHEVHECECRGLRPGMADI	246
Qy	241	QYSIAIOMMHHQPKKEICALVGFCEVCEVEMQOTLVPAKTSKXVLPALAEVPIKHE	300
Db	247	QYSIAIOMMHHQPKKEICALVGFCEVCEVEMQOTLVPAKTSKXVLPALAEVPIKHE	306
Qy	301	VPAKSVDYCEVCEFLVEVYTKLIDNNKTEKEILDAFDKMSCKLPKSLSEBQCEVVDIYGS	360
Db	307	VPAKSVDYCEVCEFLVEVYTKLIDNNKTEKEILDAFDKMSCKLPKSLSEBQCEVVDIYGS	366
Qy	361	SISILILEEVSPELVCSMLHCSGRLPALTYHTQBPQDGFCEVCKLVGYLDRLNLEKN	420
Db	367	SISILILEEVSPELVCSMLHCSGRLPALTYHTQBPQDGFCEVCKLVGYLDRLNLEKN	426
Qy	421	STJOEIIIALEKGCSEFLPDYQKQCDQFAVEAYEPVLIEVLEWMDPSFVCKIAGCPSAH	480
Db	427	STJOEIIIALEKGCSEFLPDYQKQCDQFAVEAYEPVLIEVLEWMDPSFVCKIAGCPSAH	486
Qy	481	KPLIGTEKCIWGSYVCONTEPAAOCNAVEHCKKHWN	518
Db	487	KPLIGTEKCIWGSYVCONTEPAAOCNAVEHCKKHWN	524

RESULT	10
ID	ABU05202
XX	ABU05202 standard; protein; 524 AA.
AC	ABU05202;
DT	29-JAN-2003 (first entry)
DE	Human expressed protein tag (EPT) #1868.
KW	Translational profiling; expressed protein tag; EPT; kinase, phosphatase,
KW	protease; protease inhibitor; transporter; cytoskeletal protein;
KW	receptor; transcription factor; cancer; MHC;
KW	major histocompatibility complex; myeloma; colon cancer; gastric cancer;
XX	adenocarcinoma; sarcoma; melanoma; lymphoma; leukaemia.
OS	Homo sapiens.
PN	WO200278524-A2.
PD	10-OCT-2002.
PF	28-MAR-2002; 2002WO-US009671.

XX 28-MAR-2001; 2001US-0279495P  
PR 21-MAY-2001; 2001US-0292544P  
PR 08-AUG-2001; 2001US-0310801P  
PR 01-OCT-2001; 2001US-0326370P  
PR 04-DEC-2001; 2001US-0336780P  
PR 20-FEB-2002; 2002US-0358985P  
VV

PA (ZYCO-) ZYCOS INC.  
XX  
XX  
PI Chiciz RM, Tomlinson AJ, Urban RG  
XX  
XX  
DR WPI: 2003-040607/03.  
VV

PT New polypeptides (e.g. kinases, phosphatases, proteases, transporters, cytoskeletal proteins, receptors or transcription factors), useful for treating cancer, e.g. colon cancer, gastric cancer, sarcoma, lymphoma or leukemia.

PS Example 2; SEQ ID NO 1868; 134pp; English  
XX

The invention describes a purified polypeptide, which comprises a fragment of a kinase, phosphatase, protease, or protease inhibitor, transporter, cytoskeletal protein, receptor or transcription factor. The polypeptide is useful as an immunogenic composition for eliciting in a mammal an immunogenic response directed against any of the purified polypeptide. The purified polypeptide, or the antibody that binds to this polypeptide, is useful for treating cancer. The polypeptide is also useful for identifying compounds that binds to a naturally processed class I or class II MHC-binding polypeptide. The polypeptides and polynucleotides are particularly useful for treating or preventing myeloma, colon cancer, gastric cancer, adenocarcinoma, sarcoma, melanoma, lymphoma or leukemia. These are also useful for screening agents for treating the above mentioned diseases. This sequence represents an expressed protein tag (EPT) isolated from human tissue for translational profiling. Note: This sequence does not appear in the printed specification but was obtained in electronic format directly from WIP0 at [ftp.wipo.int/pub/published\\_pct\\_sequences](http://wipo.int/pub/published_pct_sequences)

Query Match	100.0%;	Score 2759;	DB 6;	Length 524;
Best Local Similarity	100.0%;	Pred. No. 7,2e-235;		
Matches 518; Conservative	0;	Mismatches 0;	Indels 0;	Gaps 0;

QY	1	IASLILGALAGPVLGIXECTRGSAAWQCNVKTASQCGAVHICLQTVNMRPVTXSLPCD IC	60
Db	7	IASLILGALAGPVLGIXECTRGSAAWQCNVKTASQCGAVHICLQTVNMRPVTXSLPCD IC	66
QY	61	KDVTYTAAGDMLKDNATTEERILVYLEKTCMDLPRPMNASCKEIVDSYLPVLIIIKGEMS	120
Db	67	KDVTYTAAGDMLKDNATTEERILVYLEKTCMDLPRPMNASCKEIVDSYLPVLIIIKGEMS	126
QY	121	RPEVCSANLCESIQKHAEINHQKQESNKPILBMTBEVAPPMANIPLLIYPODGR	180
Db	127	RPEVCSANLCESIQKHAEINHQKQESNKPILBMTBEVAPPMANIPLLIYPODGR	186
QY	181	SKPQPKNDGVQCDDIQMTVDIQTVRINSTFYVALYEHVYEECDRIQGPMAOICKNYIS	240
Db	187	SKPQPKNDGVQCDDIQMTVDIQTVRINSTFYVALYEHVYEECDRIQGPMAOICKNYIS	246
QY	241	QYSEIALQMMHMQPKIECALVGFCEVKEPMQTLVPAKASKNVIPALEIVEPIKGE	300
Db	247	QYSEIALQMMHMQPKIECALVGFCEVKEPMQTLVPAKASKNVIPALEIVEPIKGE	306
QY	301	VAKSDVYCEVCEPIVKEVTKLIDNNKTEKSIDLAFPMCKSKLPKSLSEECQEVVDYGS	360
Db	307	VAKSDVYCEVCEPIVKEVTKLIDNNKTEKSIDLAFPMCKSKLPKSLSEECQEVVDYGS	366
QY	361	STLSLILEEVSPELVCSMLHCSGTRPLPALVHTHTQPKDGFCEVCUKLVGYIDRNLKKN	420
Db	367	STLSLILEEVSPELVCSMLHCSGTRPLPALVHTHTQPKDGFCEVCUKLVGYIDRNLKKN	426

QY 421 STKQIILAALEKGCSEFLPDPYOKCDQFAEYEPVLIILVEWMPDSFVCLKIGACPSAH 480  
 DB 427 STKQIILAALEKGCSEFLPDPYOKCDQFAEYEPVLIILVEWMPDSFVCLKIGACPSAH 486  
 QY 481 KPLIGTEKCIWGPSYWCNTETTAACNAVEHCKRHVMN 518  
 DB 487 KPLIGTEKCIWGPSYWCNTETTAACNAVEHCKRHVMN 524

RESULT 11  
 ID ABU05214 standard; protein; 524 AA.  
 AC ABU05214;  
 XX 29-JAN-2003 (first entry)  
 DT Human expressed protein tag (EPT) #1880.  
 DE  
 XX Translational profiling; expressed protein tag; EPT; kinase; phosphatase;  
 KW protease; protease inhibitor; transporter; cytoskeletal protein;  
 KW receptor; transcription factor; cancer; MHC;  
 KW major histocompatibility complex; myeloma; colon cancer; gastric cancer;  
 KW adenocarcinoma; sarcoma; melanoma; lymphoma; leukaemia.  
 XX Homo sapiens.  
 OS  
 XX WO200278524-A2.  
 EN  
 XX 10-OCT-2002.  
 PD  
 XX 28-MAR-2002; 2002MO-US009671.  
 PE  
 XX 28-MAR-2001; 2001US-0279495P.  
 PR 21-MAY-2001; 2001US-0292544P.  
 PR 08-AUG-2001; 2001US-0310801P.  
 PR 01-OCT-2001; 2001US-0326370P.  
 PR 04-DEC-2001; 2001US-0336780P.  
 PR 20-FEB-2002; 2002US-0358985P.  
 XX  
 PA (ZYCO-) ZYCOS INC.  
 XX  
 PI Chicz RM, Tomlinson AJ, Urban RG;  
 XX  
 DR WPI; 2003-040607/03.  
 XX  
 PT New polypeptides (e.g. kinases, phosphatases, proteases, transporters,  
 PT cytoskeletal proteins, receptors or transcription factors), useful for  
 PT treating cancer, e.g. colon cancer, gastric cancer, sarcoma, lymphoma or  
 PT leukemia.  
 XX  
 PS Example 2; SEQ ID NO 1880; 134p; English.  
 XX  
 CC The invention describes a purified polypeptide, which comprises a  
 CC fragment of a kinase, phosphatase, protease, protease inhibitor  
 CC transporter, cytoskeletal protein, receptor or transcription factor. The  
 CC polypeptide is useful as an immunogenic composition for eliciting in a  
 CC mammal an immunogenic response directed against any of the purified  
 CC polypeptide. The purified polypeptide, or the antibody that binds to this  
 CC polypeptide, is useful for treating cancer. The polypeptide is also  
 CC useful for identifying compounds that binds to a naturally processed  
 CC class I or class II MHC-binding polypeptide. The polypeptides and  
 CC polynucleotides are particularly useful for treating or preventing  
 CC myeloma, colon cancer, gastric cancer, adenocarcinoma, sarcoma, melanoma,  
 CC lymphoma or leukaemia. These are also useful for screening agents for  
 CC treating the above mentioned diseases. This sequence represents an  
 CC expressed protein tag (EPT) isolated from human tissue for translational  
 CC profiling. Note: This sequence does not appear in the printed  
 CC specification but was obtained in electronic format directly from WIPO at  
 CC ftp.wipo.int/pub/published\_pct\_sequences  
 XX  
 SQ Sequence 524 AA;

Query Match 100.0%; Score 2759; DB 6; Length 524;  
 Best Local Similarity 100.0%; Pred. No. 7, 2e-235;  
 Matches 518; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LASLIGALAPVIGLKECTRGSAVWQNVKTASDGA VKIQLQTVMNKPVYSIPCIC 60  
 DB 7 LASLIGALAPVIGLKECTRGSAVWQNVKTASDGA VKIQLQTVMNKPVYSIPCIC 66  
 QY 61 KDVTAAADMDLKDNTAEELIIVLEKTCMDLPKPMASCKEYDUSYLPVLIIDIIKGSMS 120  
 DB 67 KDVTAAADMDLKDNTAEELIIVLEKTCMDLPKPMASCKEYDUSYLPVLIIDIIKGSMS 126  
 QY 121 RPEVCSALNCEISIQKILAEINHQKLESNKIPELMDTEVAPPMANIPLLTYQDGP 180  
 DB 127 RPEVCSALNCEISIQKILAEINHQKLESNKIPELMDTEVAPPMANIPLLTYQDGP 186  
 QY 181 SKPQKNDGVQCQDCIQWVTDIQTAVRTNSTFVQALVEHVKEBCDRLGPMADICKNYIS 240  
 DB 187 SKPQKNDGVQCQDCIQWVTDIQTAVRTNSTFVQALVEHVKEBCDRLGPMADICKNYIS 246  
 QY 241 QYSEIATQMMHMQPKETCALVGFCDVEKEMPMQTLVPAKASKNVITPALEIYERIKGE 300  
 DB 247 QYSEIATQMMHMQPKETCALVGFCDVEKEMPMQTLVPAKASKNVITPALEIYERIKGE 306  
 QY 301 VPAKSDVYCEVCEFLVKEVTKLIDNNKTEKILDAFDGKCSKLPKSLSEECOEVDVTYGS 360  
 DB 307 VPAKSDVYCEVCEFLVKEVTKLIDNNKTEKILDAFDGKCSKLPKSLSEECOEVDVTYGS 366  
 QY 361 SILSLILEVSPBELVCSMLHLCSGTRLPALTIVHTQPRDGGCEVCKLVGLDRLNLEKN 420  
 DB 367 SILSLILEVSPBELVCSMLHLCSGTRLPALTIVHTQPRDGGCEVCKLVGLDRLNLEKN 426  
 QY 421 STKQIILAALEKGCSEFLPDPYOKCDQFAEYEPVLIILVEWMPDSFVCLKIGACPSAH 480  
 DB 427 STKQIILAALEKGCSEFLPDPYOKCDQFAEYEPVLIILVEWMPDSFVCLKIGACPSAH 486  
 QY 481 KPLIGTEKCIWGPSYWCNTETTAACNAVEHCKRHVMN 518  
 DB 487 KPLIGTEKCIWGPSYWCNTETTAACNAVEHCKRHVMN 524

RESULT 12  
 ID ABU05215 standard; protein; 524 AA.  
 AC ABU05215;  
 XX 29-JAN-2003 (first entry)  
 DT Human expressed protein tag (EPT) #1881.  
 DE  
 XX Translational profiling; expressed protein tag; EPT; kinase; phosphatase;  
 KW protease; protease inhibitor; transporter; cytoskeletal protein;  
 KW receptor; transcription factor; cancer; MHC;  
 KW major histocompatibility complex; myeloma; colon cancer; gastric cancer;  
 KW adenocarcinoma; sarcoma; melanoma; lymphoma; leukaemia.  
 XX Homo sapiens.  
 OS  
 XX WO200278524-A2.  
 EN  
 XX 10-OCT-2002.  
 PD  
 XX 28-MAR-2002; 2002MO-US009671.  
 PE  
 XX 28-MAR-2001; 2001US-0279495P.  
 PR 21-MAY-2001; 2001US-0292544P.  
 PR 08-AUG-2001; 2001US-0310801P.  
 PR 01-OCT-2001; 2001US-0326370P.  
 PR 04-DEC-2001; 2001US-0336780P.  
 PR 20-FEB-2002; 2002US-0358985P.  
 XX  
 PA (ZYCO-) ZYCOS INC.

XX Chicz RM, Tomlinson AJ, Urban RG;  
XX  
XX WPI, 2003-040607/03.  
DR

XX New polypeptides (e.g. kinases, phosphatases, proteases, transporters,  
XX cytoskeletal proteins, receptors or transcription factors), useful for  
XX treating cancer, e.g. colon cancer, gastric cancer, sarcoma, lymphoma or  
XX leukemia.

PS Example 2; SEQ ID NO 1861; 134pp; English.

XX The invention describes a purified polypeptide, which comprises a  
XX fragment of a kinase, phosphatase, protease, protease inhibitor,  
XX transporter, cytoskeletal protein, receptor or transcription factor. The  
XX polypeptide is useful as an immunogenic composition for eliciting in a  
XX mammal an immunogenic response directed against any of the purified  
XX polypeptide. The purified polypeptide, or the antibody that binds to this  
XX polypeptide, is useful for treating cancer. The polypeptide is also  
XX useful for identifying compounds that binds to a naturally processed  
XX class I or class II MHC-binding polypeptide. The polypeptide is also  
XX polynucleotides are particularly useful for treating or preventing  
XX myeloma, colon cancer, gastric cancer, adenocarcinoma, sarcoma, melanoma,  
XX lymphoma or leukemia. These are also useful for screening agents for  
XX treating the above mentioned diseases. This sequence represents an  
XX expressed protein tag (EPT) isolated from human tissue for translational  
XX profiling. Note: This sequence does not appear in the printed  
XX specification but was obtained in electronic format directly from WIPO at  
XX ftp.wipo.int/pub/published\_pct\_sequences

XX Sequence 524 AA;

Query Match 100.0%; Score 2759; DB 6; Length 524;  
Best Local Similarity 100.0%; Pred. No. 7.2e-235;

Matches 518; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY 1 LASLIGALAGPVILGKECTRGSAVWCQNVKTAASDCGAVGHCQTQVWNNKPTVSLPCDLC 60
DB 7 LASLIGALAGPVILGKECTRGSAVWCQNVKTAASDCGAVGHCQTQVWNNKPTVSLPCDLC 66
QY 61 KDVVTAAGDMLKDNATBEELIVYLEKTCQWLPRPNMSASCKEIVDSYLPVILDIIGEMS 120
DB 67 KDVVTAAGDMLKDNATBEELIVYLEKTCQWLPRPNMSASCKEIVDSYLPVILDIIGEMS 126
QY 121 RGEVCSALNLCESLQKHLAEINHOKEENKIPBLDTEVVAAPMANIPILLYPDGPR 180
DB 127 RGEVCSALNLCESLQKHLAEINHOKEENKIPBLDTEVVAAPMANIPILLYPDGPR 186
QY 181 SKPQKNDGVCCQDCIQWVTDIQTAVRTNSTFVQALVHVKEECDRIGPGWADICRNYIS 240
DB 187 SKPQKNDGVCCQDCIQWVTDIQTAVRTNSTFVQALVHVKEECDRIGPGWADICRNYIS 246
QY 241 QYSEAIQMMHMPKEICAVGFCDEYKEMPQTLVPAYKASNVIPALEVEPIKHE 300
DB 247 QYSEAIQMMHMPKEICAVGFCDEYKEMPQTLVPAYKASNVIPALEVEPIKHE 306
QY 301 VPAKSDVYCEVCEFLVKEVTKLINNKTEKEIILDAFPMKSKLPKSLSEECQEVPTGYS 360
DB 307 VPAKSDVYCEVCEFLVKEVTKLINNKTEKEIILDAFPMKSKLPKSLSEECQEVPTGYS 366
QY 361 SIISILIEEVSPELVCSMLHCSGSTRLPALTVAHTQPKDGFCEVCKLGVGLDRMLEKN 420
DB 367 SIISILIEEVSPELVCSMLHCSGSTRLPALTVAHTQPKDGFCEVCKLGVGLDRMLEKN 426
QY 421 STKOELIALAEEKGSPFDDPKOCQDOFVAEVEVULLEIVWMDSPFVCLKIGACPSAH 480
DB 427 STKOELIALAEEKGSPFDDPKOCQDOFVAEVEVULLEIVWMDSPFVCLKIGACPSAH 486
QY 481 KPLILGTEKCIWGSYWCQNTETAACQNAVEHCKRHWAN 518
DB 487 KPLILGTEKCIWGSYWCQNTETAACQNAVEHCKRHWAN 524

```

RESULT 13  
ABU05199  
ID ABU05199 standard; protein; 524 AA.  
XX

AC ABU05199;

DI 29-JAN-2003 (first entry)

DE Human expressed protein tag (EPT) #1865.

KW Translational profiling; expressed protein tag; EPT; kinase; phosphatase;  
KW protease; protease inhibitor; transporter; cytoskeletal protein;  
KW receptor; transcription factor; cancer; MHC;  
KW major histocompatibility complex; myeloma; colon cancer; gastric cancer;  
KW adenocarcinoma; sarcoma; melanoma; lymphoma; leukemia.

OS Homo sapiens.

PN WO200278524-A2.

PD 10-OCT-2002.

PF 28-MAR-2002; 2002MO-US009671.

PR 28-MAR-2001; 2001US-0279495P.

PR 21-MAY-2001; 2001US-0292544P.

PR 08-AUG-2001; 2001US-0310801P.

PR 01-OCT-2001; 2001US-0326370P.

PR 04-DEC-2001; 2001US-0336780P.

PR 20-FEB-2002; 2002US-0358985P.

PA (ZYCO-) ZYCO INC.

PI Chicz RM, Tomlinson AJ, Urban RG;

DR WPI, 2003-040607/03.

XX New polypeptides (e.g. kinases, phosphatases, proteases, transporters,  
XX cytoskeletal proteins, receptors or transcription factors), useful for  
XX treating cancer, e.g. colon cancer, gastric cancer, sarcoma, lymphoma or  
XX leukemia.

XX Example 2; SEQ ID NO 1865; 134pp; English.

XX The invention describes a purified polypeptide, which comprises a  
XX fragment of a kinase, phosphatase, protease, protease inhibitor,  
XX transporter, cytoskeletal protein, receptor or transcription factor. The  
XX polypeptide is useful as an immunogenic composition for eliciting in a  
XX mammal an immunogenic response directed against any of the purified  
XX polypeptide. The purified polypeptide, or the antibody that binds to this  
XX polypeptide, is useful for treating cancer. The polypeptide is also  
XX useful for identifying compounds that binds to a naturally processed  
XX class I or class II MHC-binding polypeptide. The polypeptide is also  
XX polynucleotides are particularly useful for treating or preventing  
XX myeloma, colon cancer, gastric cancer, adenocarcinoma, sarcoma, melanoma,  
XX lymphoma or leukemia. These are also useful for screening agents for  
XX treating the above mentioned diseases. This sequence represents an  
XX expressed protein tag (EPT) isolated from human tissue for translational  
XX profiling. Note: This sequence does not appear in the printed  
XX specification but was obtained in electronic format directly from WIPO at  
XX ftp.wipo.int/pub/published\_pct\_sequences

XX Sequence 524 AA;

Query Match 100.0%; Score 2759; DB 6; Length 524;  
Best Local Similarity 100.0%; Pred. No. 7.2e-235;

Matches 518; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY 1 LASLIGALAGPVILGKECTRGSAVWCQNVKTAASDCGAVGHCQTQVWNNKPTVSLPCDLC 60
DB 7 LASLIGALAGPVILGKECTRGSAVWCQNVKTAASDCGAVGHCQTQVWNNKPTVSLPCDLC 66
QY 61 KDVVTAAGDMLKDNATBEELIVYLEKTCQWLPRPNMSASCKEIVDSYLPVILDIIGEMS 120

```

```

Db      67 KDVTAAAGDMLKDNATBEELIIVLYLEKTCQWMLPKPMNSASCKEIVDSYLPVLLIDIKGMS 126
Qy      121 RRGVCSALNLCESIQKHLAEINHOKELESNKIPELDTEVVAAPPMANIPLLIYPODGR 180
Db      127 RRGVCSALNLCESIQKHLAEINHOKELESNKIPELDTEVVAAPPMANIPLLIYPODGR 186
Qy      181 SKPOKNDGVCQDCIQWVTDIQTAVRTNSTFVQALVHVKEECRLGPMADICKNYIS 240
Db      187 SKPOKNDGVCQDCIQWVTDIQTAVRTNSTFVQALVHVKEECRLGPMADICKNYIS 246
Qy      241 QYSEIAIQMMHMQKEICALVGFCDVEKEMQMOTLVPAKVASKNVIPAELVEPIKKE 300
Db      247 QYSEIAIQMMHMQKEICALVGFCDVEKEMQMOTLVPAKVASKNVIPAELVEPIKKE 306
Qy      301 VPAKSDVYCEVCEFLVKEVTYKLIIDNNKTEKEILDAFDKMSCKLPKSLSECOEVDITYGS 360
Db      307 VPAKSDVYCEVCEFLVKEVTYKLIIDNNKTEKEILDAFDKMSCKLPKSLSECOEVDITYGS 366
Qy      361 SILSTILLEEVSPBELVCSMLHLCSGTRLPALTVHVTQPKDGGFCVCKKLVGLDRLNLEKN 420
Db      367 SILSTILLEEVSPBELVCSMLHLCSGTRLPALTVHVTQPKDGGFCVCKKLVGLDRLNLEKN 426
Qy      421 STKOEIIAALBKSGSFLPDPYQKQCDQFAVEYEPVLIILVEVMDPSFVCLIKIGACPSAH 480
Db      427 STKOEIIAALBKSGSFLPDPYQKQCDQFAVEYEPVLIILVEVMDPSFVCLIKIGACPSAH 486
Qy      481 KPILGTEKCIWGPSYWCNTETTAACNAVEHCKRHVNN 518
Db      487 KPILGTEKCIWGPSYWCNTETTAACNAVEHCKRHVNN 524

```

## RESULT 14

ID ABU05212 standard; protein; 524 AA.

AC ABU05212;

DT 29-JAN-2003 (first entry)

DE Human expressed protein tag (EPT) #1878.

KW Translational profiling; expressed protein tag; EPT; kinase; phosphatase;

KW protease; protease inhibitor; transporter; cytoskeletal protein;

KW receptor; transcription factor; cancer; MHC;

KW major histocompatibility complex; myeloma; colon cancer; gastric cancer;

KW adenocarcinoma; sarcoma; melanoma; lymphoma; leukemia.

OS Homo sapiens.

PN WO200278524-A2.

PD 10-OCT-2002.

PF 28-MAR-2002; 2002WO-US009671.

PR 28-MAR-2001; 2001US-0278495P.

PR 21-MAY-2001; 2001US-0292544P.

PR 08-AUG-2001; 2001US-0310801P.

PR 01-OCT-2001; 2001US-0326370P.

PR 04-DEC-2001; 2001US-0336780P.

PR 20-FEB-2002; 2002US-0358955P.

PA (ZYCO-) ZYCO INC.

PI Chiciz RM, Tomlinson AJ, Urban RG;

DR WPI, 2003-040607/03.

```

XX      XX
PS      Example 2; SEQ ID NO 1878; 134pp; English.
CC      The invention describes a purified polypeptide, which comprises a
CC      fragment of a kinase, phosphatase, protease, protease inhibitor,
CC      transporter, cytoskeletal protein, receptor or transcription factor. The
CC      polypeptide is useful as an immunogenic composition for eliciting in a
CC      mammal an immunogenic response directed against any of the purified
CC      polypeptide. The purified polypeptide, or the antibody that binds to this
CC      polypeptide, is useful for treating cancer. The polypeptide is also
CC      useful for identifying compounds that binds to a naturally processed
CC      class I or class II MHC-binding polypeptide. The polypeptides and
CC      polynucleotides are particularly useful for treating or preventing
CC      myeloma, colon cancer, gastric cancer, adenocarcinoma, sarcoma, melanoma,
CC      lymphoma or leukemia. These are also useful for screening agents for
CC      treating the above mentioned diseases. This sequence represents an
CC      expressed protein tag (EPT) isolated from human tissue for translational
CC      profiling. Note: This sequence does not appear in the printed
CC      specification but was obtained in electronic format directly from WIP0 at
CC      ftp.wipo.int/pub/published_pct_sequences
CC      Sequence 524 AA:
SQ
Query Match      100.0%; Score 2759; DB 6; Length 524;
Best Local Similarity 100.0%; Pred. No. 7,2e-235;
Matches 518; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
1 LASLIGAAIAGVIGLKECTRGSAVWQNVKTASDCAVKHCLQTVNNKPTVKSIPCDIC 60
7 LASLIGAAIAGVIGLKECTRGSAVWQNVKTASDCAVKHCLQTVNNKPTVKSIPCDIC 66
Qy      61 KDVTAAAGDMLKDNATBEELIIVLYLEKTCQWMLPKPMNSASCKEIVDSYLPVLLIDIKGMS 120
Db      67 KDVTAAAGDMLKDNATBEELIIVLYLEKTCQWMLPKPMNSASCKEIVDSYLPVLLIDIKGMS 126
Qy      121 RRGVCSALNLCESIQKHLAEINHOKELESNKIPELDTEVVAAPPMANIPLLIYPODGR 180
Db      127 RRGVCSALNLCESIQKHLAEINHOKELESNKIPELDTEVVAAPPMANIPLLIYPODGR 186
Qy      181 SKPOKNDGVCQDCIQWVTDIQTAVRTNSTFVQALVHVKEECRLGPMADICKNYIS 240
Db      187 SKPOKNDGVCQDCIQWVTDIQTAVRTNSTFVQALVHVKEECRLGPMADICKNYIS 246
Qy      241 QYSEIAIQMMHMQKEICALVGFCDVEKEMQMOTLVPAKVASKNVIPAELVEPIKKE 300
Db      247 QYSEIAIQMMHMQKEICALVGFCDVEKEMQMOTLVPAKVASKNVIPAELVEPIKKE 306
Qy      301 VPAKSDVYCEVCEFLVKEVTYKLIIDNNKTEKEILDAFDKMSCKLPKSLSECOEVDITYGS 360
Db      307 VPAKSDVYCEVCEFLVKEVTYKLIIDNNKTEKEILDAFDKMSCKLPKSLSECOEVDITYGS 366
Qy      361 SILSTILLEEVSPBELVCSMLHLCSGTRLPALTVHVTQPKDGGFCVCKKLVGLDRLNLEKN 420
Db      367 SILSTILLEEVSPBELVCSMLHLCSGTRLPALTVHVTQPKDGGFCVCKKLVGLDRLNLEKN 426
Qy      421 STKOEIIAALBKSGSFLPDPYQKQCDQFAVEYEPVLIILVEVMDPSFVCLIKIGACPSAH 480
Db      427 STKOEIIAALBKSGSFLPDPYQKQCDQFAVEYEPVLIILVEVMDPSFVCLIKIGACPSAH 486
Qy      481 KPILGTEKCIWGPSYWCNTETTAACNAVEHCKRHVNN 518
Db      487 KPILGTEKCIWGPSYWCNTETTAACNAVEHCKRHVNN 524

```

## RESULT 15

ID ABU05213 standard; protein; 524 AA.

AC ABU05213;

DT 29-JAN-2003 (first entry)

DE Human expressed protein tag (EPT) #1879.

XX Translational profiling: expressed protein tag; EPT; kinase; phosphatase;  
KM protease; protease inhibitor; transporter; cytoskeletal protein;  
KM receptor; transcription factor; cancer; MHC;  
KM major histocompatibility complex; myeloma; colon cancer; gastric cancer;  
KM adenocarcinoma; sarcoma; melanoma; lymphoma; leukemia.  
OS Homo sapiens.  
XX  
XX W0200278524-A2.  
PN  
PD 10-OCT-2002.  
PF  
PF 28-MAR-2002; 2002MO-US009671.  
XX  
PR 28-MAR-2001; 2001US-0279495P.  
PR 21-MAY-2001; 2001US-0292544P.  
PR 08-AUG-2001; 2001US-0310801P.  
PR 01-OCT-2001; 2001US-0326370P.  
PR 04-DEC-2001; 2001US-0336780P.  
PR 20-FEB-2002; 2002US-0358985P.  
XX  
XX (ZYCO-) ZYCO INC.  
XX  
XX Chicz RM, Tomlinson AJ, Urban RG;  
PI  
PI MPI; 2003-040607/03.  
DR  
XX  
XX  
XX New polypeptides (e.g. kinases, phosphatases, proteases, transporters,  
PT cytoskeletal proteins, receptors or transcription factors), useful for  
PT treating cancer, e.g. colon cancer, gastric cancer, sarcoma, lymphoma or  
PT leukemia.  
XX  
XX  
XX Example 2; SEQ ID NO 1879; 134p; English.  
XX  
XX The invention describes a purified polypeptide, which comprises a  
CC fragment of a kinase, phosphatase, protease, protease inhibitor,  
CC transporter, cytoskeletal protein, receptor or transcription factor. The  
CC polypeptide is useful as an immunogenic composition for eliciting in a  
CC mammal an immunogenic response directed against any of the purified  
CC polypeptide. The purified polypeptide, or the antibody that binds to this  
CC polypeptide, is useful for treating cancer. The polypeptide is also  
CC useful for identifying compounds that binds to a naturally processed  
CC class I or class II MHC-binding polypeptide. The polypeptides and  
CC polynucleotides are particularly useful for treating or preventing  
CC myeloma, colon cancer, gastric cancer, adenocarcinoma, sarcoma, melanoma,  
CC lymphoma or leukemia. These are also useful for screening agents for  
CC treating the above mentioned diseases. This sequence represents an  
CC expressed protein tag (EPT) isolated from human tissue for translational  
CC profiling. Note: This sequence does not appear in the printed  
CC specification but was obtained in electronic format directly from WIPO at  
CC ftp.wipo.int/pub/published\_pct\_sequences  
XX  
XX Sequence 524 AA;  
SQ

Query Match 100.0%; Score 2759; DB 6; Length 524;  
Best Local Similarity 100.0%; Pred. No. 7,2e-235;  
Matches 518; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 LASLLGAAALAPVGLKECTRGSAVWCQNVKASDCAVHCLQTVWNRKPTVKSLPDIDIC 60  
Db 7 LASLLGAAALAPVGLKECTRGSAVWCQNVKASDCAVHCLQTVWNRKPTVKSLPDIDIC 66  
Qy 61 KDVTYAGDMKDNATEEELIVLEKTCDMLEPKNMSASCKEIVDSYLPYILDIIGKEMS 120  
Db 67 KDVTYAGDMKDNATEEELIVLEKTCDMLEPKNMSASCKEIVDSYLPYILDIIGKEMS 126  
Qy 121 RFGVCSALNLCESLQGHAEINHQKQESNKKIPELDMTEVAPFMANIFPLILYPODGPR 180  
Db 127 RFGVCSALNLCESLQGHAEINHQKQESNKKIPELDMTEVAPFMANIFPLILYPODGPR 186  
Qy 181 SKPOKNDGVDVCCDICIQWVTDIQTAVRTNSTFVALVHVEKCDRLPGMADICKNYIS 240  
Db 187 SKPOKNDGVDVCCDICIQWVTDIQTAVRTNSTFVALVHVEKCDRLPGMADICKNYIS 246

Qy 241 QYSEIAIOMMMHMQPKKICALVGFCDVEKEMPMQTLVPAYASKNVTPALELVEPIKGE 300  
Db 247 QYSEIAIOMMMHMQPKKICALVGFCDVEKEMPMQTLVPAYASKNVTPALELVEPIKGE 306  
Qy 301 VPAKSDVYCEVCEFLVKEVTLLIDNNKTEKEILDAFDMCSKLPKSLSECCQEVDPYGS 360  
Db 307 VPAKSDVYCEVCEFLVKEVTLLIDNNKTEKEILDAFDMCSKLPKSLSECCQEVDPYGS 366  
Qy 361 SILSILIEVSPPELVCSMLHCSGTRLPALTTHVTOPKDGFCFECVKKLGVYLDRLNLEKN 420  
Db 367 SILSILIEVSPPELVCSMLHCSGTRLPALTTHVTOPKDGFCFECVKKLGVYLDRLNLEKN 426  
Qy 421 STKOEIIAALKEKGSFLPDPYOKOCDFVAEYEPVLEILVEYMDSPFVCLKTGACPSAH 480  
Db 427 STKOEIIAALKEKGSFLPDPYOKOCDFVAEYEPVLEILVEYMDSPFVCLKTGACPSAH 486  
Qy 481 KPLLGTCKIWMGPRYWCNTETAAQCNAVEHCKRRHVN 518  
Db 487 KPLLGTCKIWMGPRYWCNTETAAQCNAVEHCKRRHVN 524

Search completed: May 5, 2004, 13:31:04  
Job time : 73.3134 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

## OM protein - protein search, using sw model

Run on: May 5, 2004, 13:10:37 ; Search time 22.6287 Seconds

(without alignments)  
1181.786 Million cell updates/sec

Title: US-09-743-684a-1\_COPY\_7\_524

Perfect score: 2759

Sequence: 1 LASLIGALAPVGLKECT.....NTETAAQCNAYEHCKRHWN 518

Scoring table: BLOSUM62

Gapop 10.0, Gapext 0.5

Searched: 389414 seqs, 51625971 residues

Total number of hits satisfying chosen parameters: 389414

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database: Issued Patents AA:\*

1: /cgm2\_6/ptodata/2/1aa/5A\_COMB.pep:\*  
2: /cgm2\_6/ptodata/2/1aa/5B\_COMB.pep:\*  
3: /cgm2\_6/ptodata/2/1aa/6A\_COMB.pep:\*  
4: /cgm2\_6/ptodata/2/1aa/6B\_COMB.pep:\*  
5: /cgm2\_6/ptodata/2/1aa/PCTUS\_COMB.pep:\*  
6: /cgm2\_6/ptodata/2/1aa/backfile1.pep:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2759	100.0	524	US-09-352-548-1	Sequence 1, Appl1
2	2736.5	99.2	523	US-08-100-247-2	Sequence 2, Appl1
3	2736.5	99.2	523	US-08-756-031-2	Sequence 2, Appl1
4	2729.5	98.9	523	US-08-232-513A-3	Sequence 3, Appl1
5	2727.5	98.9	523	US-08-483-146A-2	Sequence 2, Appl1
6	2727.5	98.9	523	US-08-484-594A-2	Sequence 2, Appl1
7	2727.5	98.9	523	US-09-076-258A-2	Sequence 2, Appl1
8	431	15.6	81	US-09-352-548-2	Sequence 2, Appl1
9	418.5	15.2	80	US-08-584-671-15	Sequence 15, Appl1
10	418.5	15.2	80	US-09-027-376-15	Sequence 15, Appl1
11	415.5	15.1	80	US-09-094-192-15	Sequence 15, Appl1
12	412	14.9	80	US-08-100-247-3	Sequence 3, Appl1
13	412	14.9	80	US-08-483-146A-3	Sequence 3, Appl1
14	412	14.9	80	US-08-232-513A-4	Sequence 4, Appl1
15	412	14.9	80	US-08-484-594A-3	Sequence 4, Appl1
16	412	14.9	80	US-09-076-258A-3	Sequence 3, Appl1
17	412	14.9	80	US-08-756-031-3	Sequence 3, Appl1
18	341.5	12.4	381	US-09-193-877-2	Sequence 3, Appl1
19	327	11.9	61	US-08-584-671-13	Sequence 13, Appl1
20	327	11.9	61	US-09-027-376-13	Sequence 13, Appl1
21	327	11.9	61	US-09-094-192-13	Sequence 13, Appl1
22	327	11.9	69	US-09-268-070-2	Sequence 13, Appl1
23	322	11.7	79	US-08-584-671-16	Sequence 16, Appl1
24	322	11.7	79	US-09-027-376-16	Sequence 16, Appl1
25	322	11.7	79	US-09-094-192-16	Sequence 16, Appl1
26	321	11.6	79	US-08-584-671-14	Sequence 14, Appl1
27	321	11.6	79	US-09-027-376-14	Sequence 14, Appl1

28	321	11.6	79	US-09-094-192-14	Sequence 14, Appl1
29	268.5	9.7	257	US-08-596-684F-7	Sequence 7, Appl1
30	239.5	8.7	60	US-08-584-671-12	Sequence 12, Appl1
31	239.5	8.7	60	US-09-027-376-12	Sequence 12, Appl1
32	239.5	8.7	60	US-09-268-070-4	Sequence 4, Appl1
33	239.5	8.7	60	US-09-094-192-12	Sequence 12, Appl1
34	238.5	8.6	68	US-09-268-070-1	Sequence 1, Appl1
35	208	7.5	67	US-09-268-070-3	Sequence 3, Appl1
36	168.5	6.1	181	US-08-848-580-12	Sequence 12, Appl1
37	168.5	6.1	181	US-08-488-123-12	Sequence 12, Appl1
38	126.5	4.6	514	US-09-107-532A-6282	Sequence 6282, Ap
39	115.5	4.2	1002	US-09-762-724-4	Sequence 4, Appl1
40	113	4.1	1027	US-09-762-724-8	Sequence 8, Appl1
41	113	4.1	1027	US-08-973-462-8	Sequence 8, Appl1
42	112.5	4.1	1014	US-09-762-724-2	Sequence 2, Appl1
43	111	4.0	22	US-08-100-247-1	Sequence 1, Appl1
44	111	4.0	22	US-08-483-146A-1	Sequence 1, Appl1
45	111	4.0	22	US-08-483-146A-11	Sequence 11, Appl1

## ALIGNMENTS

```
RESULT 1
US-09-352-548-1
; Sequence 1, Application US/09352548
; Patent No. 6500431
; GENERAL INFORMATION:
; APPLICANT: G111, Parkash S.
; TITLE OF INVENTION: No. 6500431el Inhibitors of Angiogenesis and Tumor Growth
; FILE REFERENCE: 017986-000410US
; CURRENT FILING DATE: 1999-07-12
; EARLIER APPLICATION NUMBER: US 60/092,647
; EARLIER FILING DATE: 1998-07-13
; NUMBER OF SEQ ID NOS: 59
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 1
; LENGTH: 524
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: prosaposin
; NAME/KEY: PEPTIDE
; LOCATION: (195)..(275)
; OTHER INFORMATION: Saposin B
US-09-352-548-1
Query Match 100.0%; Score 2759; DB 4; Length 524;
Best Local Similarity 100.0%; Pred. No. 9.8e-240;
Matches 518; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 LASLIGALAPVGLKECTGSAVWCNVYTAPOGAVKXICLOVWKKPTYSKPCDIC 60
Db 7 LASLIGALAPVGLKECTGSAVWCNVYTAPOGAVKXICLOVWKKPTYSKPCDIC 66
QY 61 KDVTAAADMLKDNATEEELIVYLEKTCMDLPRKPMASCKEIVDSYLPVILDIITKGENS 120
Db 67 KDVTAAADMLKDNATEEELIVYLEKTCMDLPRKPMASCKEIVDSYLPVILDIITKGENS 126
QY 121 RPEVCSALNLCESLQKHLAEINHQKLESNKIPELDTEVVAAPPMANIPLILYPQDGR 180
Db 127 RPEVCSALNLCESLQKHLAEINHQKLESNKIPELDTEVVAAPPMANIPLILYPQDGR 186
QY 181 SKPOPKNDGVCODICQWVTDIOTAVRNSFVALVHVKEBCDRLPGMADICQNTIS 240
Db 187 SKPOPKNDGVCODICQWVTDIOTAVRNSFVALVHVKEBCDRLPGMADICQNTIS 246
QY 241 QYSEIALIQMMHMQPKETICALVGFCEVKEKPMQTLVPAKYASKNVILPALEIVPIKGE 300
Db 247 QYSEIALIQMMHMQPKETICALVGFCEVKEKPMQTLVPAKYASKNVILPALEIVPIKGE 306
```



QY 301 VPAKSDVYCEVCEFLVKEVTKLIDNNKTEKEIIDAEDKMSKLPKSLSECCQEVVDYGS 360  
DB 307 VPAKSDVYCEVCEFLVKEVTKLIDNNKTEKEIIDAEDKMSKLPKSLSECCQEVVDYGS 366  
QY 361 SIISILLEEVSPELVCSMLHLCGSTRLPALTVAHTQPKDGFCEVCKLVGYLDRMLEKN 420  
DB 367 SIISILLEEVSPELVCSMLHLCGSTRLPALTVAHTQPKDGFCEVCKLVGYLDRMLEKN 426  
QY 421 STQOEIIAALKKCSFLPDYQKCDQFVAEYEPVLIIEILVEVMDPSFVCLKIGACPSAH 480  
DB 427 STQOEIIAALKKCSFLPDYQKCDQFVAEYEPVLIIEILVEVMDPSFVCLKIGACPSAH 486  
QY 481 KPILGTEKCIWGPSYWCNTETAOCNAVEHCKRHVWN 518  
DB 487 KPILGTEKCIWGPSYWCNTETAOCNAVEHCKRHVWN 524

## RESULT 2

US-08-100-247-2  
Sequence 2, Application US/08100247  
Patent No. 5571787  
GENERAL INFORMATION:  
APPLICANT: O'BRIEN, JOHN S.  
APPLICANT: KISHIMOTO, YASUO  
TITLE OF INVENTION: PROSAPOSIN AS A NEUTROTROPIC FACTOR  
NUMBER OF SEQUENCES: 5  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: KNOBBE, MARTENS, OLSON AND BEAR  
STREET: 620 NEWPORT CENTER DRIVE SIXTEENTH FLOOR  
CITY: NEWPORT BEACH  
STATE: CA  
COUNTRY: USA  
ZIP: 92660  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/100,247  
FILING DATE: 19930730  
CLASSIFICATION: 514  
ATTORNEY/AGENT INFORMATION:  
NAME: Israelien, Ned A.  
REGISTRATION NUMBER: 29,655  
REFERENCE/DOCKET NUMBER: O'Brien.002A  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 619-235-8550  
TELEFAX: 619-235-0176  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 523 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
HYPOTHEICAL: NO  
ANTI-SENSE: NO  
FRAGMENT TYPE: N-terminal  
IMMEDIATE SOURCE:  
CLONE: PROSAPOSIN  
US-08-100-247-2

Query Match 99.24; Score 2736.5; DB 1; Length 523;  
Best Local Similarity 99.44; Pred. No. 1e-237;  
Matches 515; Conservative 2; Mismatches 0; Indels 1; Gaps 1;

QY 1 LASLIGALAGPVLGKECTRGSAVWCONVKTASDDGAVHGCQTWNNKPTVSLPCDTC 60  
DB 7 LASLIGALAGPVLGKECTRGSAVWCONVKTASDDGAVHGCQTWNNKPTVSLPCDTC 66  
QY 61 KDVVTAAGDMLKDNATEEILVLYLEKTCDWLPRNMSASCKEIVDSYLPVLIIDIIIGEMS 120

DB 67 KDVVTAAGDMLKDNATEEILVLYLEKTCDWLPRNMSASCKEIVDSYLPVLIIDIIIGEMS 126  
QY 121 RPEVCSALNLCESLOHIAELNHQKOLBSNKIPDELDMTEVAVPMANIPLLLYPDQDPR 180  
DB 127 RPEVCSALNLCESLOHIAELNHQKOLBSNKIPDELDMTEVAVPMANIPLLLYPDQDPR 186  
QY 181 SKPOKNDGDVCCDCIOMVNDIOTAVRTNSFVQALVEHKEEDRLGPGMAIDCKYIIS 240  
DB 187 SKPOKNDGDVCCDCIOMVNDIOTAVRTNSFVQALVEHKEEDRLGPGMAIDCKYIIS 245  
QY 241 QYSEIAIQMMHMQPKXCALVGCDEVKEMPMQTLVPAKVASKNVLPALVELPIKHE 300  
DB 246 QYSEIAIQMMHMQPKXCALVGCDEVKEMPMQTLVPAKVASKNVLPALVELPIKHE 305  
QY 301 VPAKSDVYCEVCEFLVKEVTKLIDNNKTEKEIIDAEDKMSKLPKSLSECCQEVVDYGS 360  
DB 306 VPAKSDVYCEVCEFLVKEVTKLIDNNKTEKEIIDAEDKMSKLPKSLSECCQEVVDYGS 365  
QY 361 SIISILLEEVSPELVCSMLHLCGSTRLPALTVAHTQPKDGFCEVCKLVGYLDRMLEKN 420  
DB 366 SIISILLEEVSPELVCSMLHLCGSTRLPALTVAHTQPKDGFCEVCKLVGYLDRMLEKN 425  
QY 421 STQOEIIAALKKCSFLPDYQKCDQFVAEYEPVLIIEILVEVMDPSFVCLKIGACPSAH 480  
DB 426 STQOEIIAALKKCSFLPDYQKCDQFVAEYEPVLIIEILVEVMDPSFVCLKIGACPSAH 485  
QY 481 KPILGTEKCIWGPSYWCNTETAOCNAVEHCKRHVWN 518  
DB 486 KPILGTEKCIWGPSYWCNTETAOCNAVEHCKRHVWN 523

## RESULT 3

US-08-756-031-2  
Sequence 2, Application US/08756031  
Patent No. 6590074  
GENERAL INFORMATION:  
APPLICANT: O'BRIEN, JOHN S.  
APPLICANT: KISHIMOTO, YASUO  
TITLE OF INVENTION: PROSAPOSIN AS A NEUTROTROPIC FACTOR  
NUMBER OF SEQUENCES: 5  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: KNOBBE, MARTENS, OLSON AND BEAR  
STREET: 620 NEWPORT CENTER DRIVE SIXTEENTH FLOOR  
CITY: NEWPORT BEACH  
STATE: CA  
COUNTRY: USA  
ZIP: 92660  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/756,031  
FILING DATE: 26-NOV-1996  
CLASSIFICATION: 514  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/100,247  
FILING DATE: 30-JUL-1993  
ATTORNEY/AGENT INFORMATION:  
NAME: Israelien, Ned A.  
REGISTRATION NUMBER: 29,655  
REFERENCE/DOCKET NUMBER: O'Brien.002A  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 619-235-8550  
TELEFAX: 619-235-0176  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 523 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear



MOLECULE TYPE: peptide  
 HYPOTHETICAL: NO  
 ANTI-SENSE: NO  
 FRAGMENT TYPE: N-terminal  
 IMMEDIATE SOURCE:  
 CLONE: PROSAPOSIN  
 US-08-756-031-2

Query Match 99.2%; Score 2736.5; DB 4; Length 523;  
 Best Local Similarity 99.4%; Pred. No. 1e-237;  
 Matches 515; Conservative 2; Mismatches 0; Indels 1; Gaps 1;

QY 1 LASLIGALAGPVGLKECTRGSAVWCONVKTASDCGAVKHCLOTWNKPTVKSIPCDIC 60  
 DB 7 LASLIGALAGPVGLKECTRGSAVWCONVKTASDCGAVKHCLOTWNKPTVKSIPCDIC 66  
 QY 61 KDVTAAAGDMKDNATEEEIIVLEKTCMDLPRKPMASASCEIYDSYLPVLDIIGKEMS 120  
 DB 67 KDVTAAAGDMKDNATEEEIIVLEKTCMDLPRKPMASASCEIYDSYLPVLDIIGKEMS 126  
 QY 121 RRGVCSALNLCESIQKHLAELNHQKLESNKIPELDMTEVAVAPPMANIPLLVPQDGP 180  
 DB 127 RRGVCSALNLCESIQKHLAELNHQKLESNKIPELDMTEVAVAPPMANIPLLVPQDGP 186  
 QY 181 SKPQKNGDVQCDCIQWTDIQTAVRTNSTFFVALVEHVKEBCDRGLPGMADICKNYIS 240  
 DB 187 SKPQKNGDVQCDCIQWTDIQTAVRTNSTFFVALVEHVKEBCDRGLPGMADICKNYIS 245  
 QY 241 QYSEIATQMMHMPKEICALVGFCEVKEPMQTLVPAKASKNVLPALVPEIKKHE 300  
 DB 246 QYSEIATQMMHMPKEICALVGFCEVKEPMQTLVPAKASKNVLPALVPEIKKHE 305  
 QY 301 VPAKSDVYCEVCEFLVKEVTKLIDNNKTEKEILDAFDMCKSLPKSLSEECQEVVDITYGS 360  
 DB 306 VPAKSDVYCEVCEFLVKEVTKLIDNNKTEKEILDAFDMCKSLPKSLSEECQEVVDITYGS 365  
 QY 361 SILSILAEVSPPELVCSMLHLCSTGRPLPALTVHTQPKDGFCEVCKKLVGLDRNLEKN 420  
 DB 366 SILSILAEVSPPELVCSMLHLCSTGRPLPALTVHTQPKDGFCEVCKKLVGLDRNLEKN 425  
 QY 421 STKOEIIAALEKGSFLPDYQKOCDFVAEYEPVLLEIIVWDPSPVCLIKGACPSAH 480  
 DB 426 STKOEIIAALEKGSFLPDYQKOCDFVAEYEPVLLEIIVWDPSPVCLIKGACPSAH 485  
 QY 481 KPLGTETKCIWGPSTWCONTEETAACNAVEHCKRHVNN 518  
 DB 486 KPLGTETKCIWGPSTWCONTEETAACNAVEHCKRHVNN 523

## RESULT 4

US-08-232-513A-3  
 ; Sequence 3, Application US/08232513A  
 ; Patent No. 5700909  
 ; GENERAL INFORMATION:  
 ; APPLICANT: O'Brien, John S.  
 ; TITLE OF INVENTION: prosaposin and cytokine-Derived Peptides  
 ; TITLE OF INVENTION: as Therapeutic Agents  
 ; NUMBER OF SEQUENCES: 20  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: Campbell & Flores LLP  
 ; STREET: 4370 La Jolla Village Drive, Suite 700  
 ; CITY: San Diego  
 ; STATE: California  
 ; COUNTRY: USA  
 ; ZIP: 92122  
 ; COMPUTER READABLE FORM:  
 ; MEDIUM TYPE: floppy disk  
 ; COMPUTER: IBM PC compatible  
 ; OPERATING SYSTEM: PC-DOS/MS-DOS  
 ; SOFTWARE: Patentin Release #1.0, Version #1.25  
 ; CURRENT APPLICATION DATA:  
 ; APPLICATION NUMBER: US/08/232, 513A  
 ; FILING DATE: 21-APR-1994

CLASSIFICATION: 514  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: US 08/100,247  
 FILING DATE: 30-JUL-1993  
 ATTORNEY/AGENT INFORMATION:  
 NAME: Campbell, Cathryn A.  
 REGISTRATION NUMBER: 31,815  
 REFERENCE/DOCKET NUMBER: P-UD 1643  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: (619) 535-9001  
 TELEFAX: (619) 535-8949  
 INFORMATION FOR SEQ ID NO: 3:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 523 amino acids  
 TYPE: amino acid  
 TOPOLOGY: linear  
 MOLECULE TYPE: protein  
 FEATURE:  
 NAME/KEY: Protein  
 LOCATION: 1..523  
 OTHER INFORMATION: /label= Hum\_prosaposin  
 US-08-232-513A-3

Query Match 98.9%; Score 2729.5; DB 1; Length 523;  
 Best Local Similarity 99.2%; Pred. No. 4.4e-237;  
 Matches 514; Conservative 2; Mismatches 1; Indels 1; Gaps 1;

QY 1 LASLIGALAGPVGLKECTRGSAVWCONVKTASDCGAVKHCLOTWNKPTVKSIPCDIC 60  
 DB 7 LASLIGALAGPVGLKECTRGSAVWCONVKTASDCGAVKHCLOTWNKPTVKSIPCDIC 66  
 QY 61 KDVTAAAGDMKDNATEEEIIVLEKTCMDLPRKPMASASCEIYDSYLPVLDIIGKEMS 120  
 DB 67 KDVTAAAGDMKDNATEEEIIVLEKTCMDLPRKPMASASCEIYDSYLPVLDIIGKEMS 126  
 QY 121 RRGVCSALNLCESIQKHLAELNHQKLESNKIPELDMTEVAVAPPMANIPLLVPQDGP 180  
 DB 127 RRGVCSALNLCESIQKHLAELNHQKLESNKIPELDMTEVAVAPPMANIPLLVPQDGP 186  
 QY 181 SKPQKNGDVQCDCIQWTDIQTAVRTNSTFFVALVEHVKEBCDRGLPGMADICKNYIS 240  
 DB 187 SKPQKNGDVQCDCIQWTDIQTAVRTNSTFFVALVEHVKEBCDRGLPGMADICKNYIS 245  
 QY 241 QYSEIATQMMHMPKEICALVGFCEVKEPMQTLVPAKASKNVLPALVPEIKKHE 300  
 DB 246 QYSEIATQMMHMPKEICALVGFCEVKEPMQTLVPAKASKNVLPALVPEIKKHE 305  
 QY 301 VPAKSDVYCEVCEFLVKEVTKLIDNNKTEKEILDAFDMCKSLPKSLSEECQEVVDITYGS 360  
 DB 306 VPAKSDVYCEVCEFLVKEVTKLIDNNKTEKEILDAFDMCKSLPKSLSEECQEVVDITYGS 365  
 QY 361 SILSILAEVSPPELVCSMLHLCSTGRPLPALTVHTQPKDGFCEVCKKLVGLDRNLEKN 420  
 DB 366 SILSILAEVSPPELVCSMLHLCSTGRPLPALTVHTQPKDGFCEVCKKLVGLDRNLEKN 425  
 QY 421 STKOEIIAALEKGSFLPDYQKOCDFVAEYEPVLLEIIVWDPSPVCLIKGACPSAH 480  
 DB 426 STKOEIIAALEKGSFLPDYQKOCDFVAEYEPVLLEIIVWDPSPVCLIKGACPSAH 485  
 QY 481 KPLGTETKCIWGPSTWCONTEETAACNAVEHCKRHVNN 518  
 DB 486 KPLGTETKCIWGPSTWCONTEETAACNAVEHCKRHVNN 523

## RESULT 5

US-08-483-146A-2  
 ; Sequence 2, Application US/08483146A  
 ; Patent No. 5696080  
 ; GENERAL INFORMATION:  
 ; APPLICANT: O'Brien, John S.  
 ; APPLICANT: Kishimoto, Yasuo  
 ; TITLE OF INVENTION: PHARMACEUTICAL COMPOSITIONS  
 ; TITLE OF INVENTION: COMPRISING PROSAPOSIN AND NEUROTROPHIC PEPTIDES DERIVED

TITLE OF INVENTION: THEREFROM  
NUMBER OF SEQUENCES: 11  
CORRESPONDENCE ADDRESSES:  
ADDRESSEE: Knobbe, Martens, Olson and Bear  
STREET: 620 Newport Center Blvd. 16th Floor  
CITY: Newport Beach  
STATE: CA  
COUNTRY: USA  
ZIP: 92660  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: DOS  
SOFTWARE: FastSeq for Windows Version 2.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/483,146A  
FILING DATE: 07-JUN-1995  
CLASSIFICATION: 514  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER:  
FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: Israel, Ned A  
REGISTRATION NUMBER: 29,655  
REFERENCE/DOCKET NUMBER: MYELOS.002DV1  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 619-235-8550  
TELEFAX: 619-235-0176  
TELEX:  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 523 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
FRAGMENT TYPE: N-terminal  
US-08-483-146A-2

Query Match 98.9%; Score 2727.5; DB 1; Length 523;  
Best Local Similarity 99.2%; Pred. No. 6.7e-237;  
Matches 514; Conservative 2; Mismatches 1; Indels 1; Gaps 1;  
QY 1 LASLIGALAPVGLKECTRGSAVWCQNVKTASDCGAVKHCLOTWNKPTVKSLPEDIC 60  
DB 7 LASLIGALAPVGLKECTRGSAVWCQNVKTASDCGAVKHCLOTWNKPTVKSLPEDIC 66  
QY 61 KDVTAAAGDMKDNATEEELIVYEKTCMDLPRKPMNSASCKEIVDSYLPVILDIIGKEMS 120  
DB 67 KDVTAAAGDMKDNATEEELIVYEKTCMDLPRKPMNSASCKEIVDSYLPVILDIIGKEMS 126  
QY 121 RRGVCSALNLCESLQKHLAELNHQKLESNKIPELDMTEVVAFPMANIPLLIYPQDGR 180  
DB 127 RRGVCSALNLCESLQKHLAELNHQKLESNKIPELDMTEVVAFPMANIPLLIYPQDGR 186  
QY 181 SKPOKNGDVGCOCICQMTDIOAVRNTSTFVALVEHKEBDRIGPMADICKNYIS 240  
DB 187 SKPOKNGDVGCOCICQMTDIOAVRNTSTFVALVEHKEBDRIGPMADICKNYIS 245  
QY 241 QYSIAIOMMMHMQKEICALVGFCEVKEPMQTLVPAKASKNVPALEVEPIKKHE 300  
DB 246 QYSIAIOMMMHMQKEICALVGFCEVKEPMQTLVPAKASKNVPALEVEPIKKHE 305  
QY 301 VPAASDVYCEVCELVKEVTKLIDNNKTEKEIIDAPKMSCKLPSKISECEQEVDTYGS 360  
DB 306 VPAASDVYCEVCELVKEVTKLIDNNKTEKEIIDAPKMSCKLPSKISECEQEVDTYGS 365  
QY 361 SIISILIEVSPELVCSMLHLCSTRLPALTVHTOPKOGFGEVCKKIVGYIDRLLEN 420  
DB 366 SIISILIEVSPELVCSMLHLCSTRLPALTVHTOPKOGFGEVCKKIVGYIDRLLEN 425  
QY 421 STGOEILALBKGCSTLPDPYQKOCQFVAEYEPVLEILVEVNDPSFVCLKIGACPSAH 480

DB 426 STGOEILALBKGCSTLPDPYQKOCQFVAEYEPVLEILVEVNDPSFVCLKIGACPSAH 485  
QY 481 KPLIGTEKICWGPWYQONTTAAQCNVAHECKRHVNN 518  
DB 486 KPLIGTEKICWGPWYQONTTAAQCNVAHECKRHVNN 523  
RESULT 6  
US-08-484-594A-2  
Sequence 2, Application US/08484594A  
Patent No. 5714459  
GENERAL INFORMATION:  
APPLICANT: O'Brien, John S.  
APPLICANT: Kishimoto, Yasuo  
TITLE OF INVENTION: USE OF PROSAPOSIN AND NEUROTROPHIC PEPTIDES  
NUMBER OF SEQUENCES: 11  
CORRESPONDENCE ADDRESSES:  
ADDRESSEE: Knobbe, Martens, Olson and Bear  
STREET: 620 Newport Center Drive, Sixteenth Floor  
CITY: Newport Beach  
STATE: CA  
COUNTRY: USA  
ZIP: 92660  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: DOS  
SOFTWARE: FastSeq for Windows Version 2.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/484,594A  
FILING DATE: 07-JUN-1995  
CLASSIFICATION: 514  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/100,247  
FILING DATE: 30-JUL-1993  
ATTORNEY/AGENT INFORMATION:  
NAME: Israel, Ned A  
REGISTRATION NUMBER: 29,655  
REFERENCE/DOCKET NUMBER: MYELOS.002DV2  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 619-235-8550  
TELEFAX: 619-235-0176  
TELEX:  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 523 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
FRAGMENT TYPE: N-terminal  
US-08-484-594A-2  
Query Match 98.9%; Score 2727.5; DB 1; Length 523;  
Best Local Similarity 99.2%; Pred. No. 6.7e-237;  
Matches 514; Conservative 2; Mismatches 1; Indels 1; Gaps 1;  
QY 1 LASLIGALAPVGLKECTRGSAVWCQNVKTASDCGAVKHCLOTWNKPTVKSLPEDIC 60  
DB 7 LASLIGALAPVGLKECTRGSAVWCQNVKTASDCGAVKHCLOTWNKPTVKSLPEDIC 66  
QY 61 KDVTAAAGDMKDNATEEELIVYEKTCMDLPRKPMNSASCKEIVDSYLPVILDIIGKEMS 120  
DB 67 KDVTAAAGDMKDNATEEELIVYEKTCMDLPRKPMNSASCKEIVDSYLPVILDIIGKEMS 126  
QY 121 RRGVCSALNLCESLQKHLAELNHQKLESNKIPELDMTEVVAFPMANIPLLIYPQDGR 180  
DB 127 RRGVCSALNLCESLQKHLAELNHQKLESNKIPELDMTEVVAFPMANIPLLIYPQDGR 186  
QY 181 SKPOKNGDVGCOCICQMTDIOAVRNTSTFVALVEHKEBDRIGPMADICKNYIS 240  
DB 187 SKPOKNGDVGCOCICQMTDIOAVRNTSTFVALVEHKEBDRIGPMADICKNYIS 245

QY 241 QYSEIAIQMMHMQKEICALVGFCDVEKEMPMQTLVPAKVASKNVLPALVEPIKKE 300  
Db 246 QYSEIAIQMMHMQKEICALVGFCDVEKEMPMQTLVPAKVASKNVLPALVEPIKKE 305  
QY 301 VPAKSDVYCEVCEFLVKEVTYKLIINNKTKEKILDAFDMKCSKLPKSLSEECQEVVDITYGS 360  
Db 306 VPAKSDVYCEVCEFLVKEVTYKLIINNKTKEKILDAFDMKCSKLPKSLSEECQEVVDITYGS 365  
QY 361 SILSLILEEVSPELVCSMLHLCSTGRLPALVHTVHTOPKDGFCVCKLVGYLDRLNEKN 420  
Db 366 SILSLILEEVSPELVCSMLHLCSTGRLPALVHTVHTOPKDGFCVCKLVGYLDRLNEKN 425  
QY 421 STKQELIAALEKGSFLPDYPQKODQFVAEYEPVLIEILVEWMDPSFVCLKIGACPSAH 480  
Db 426 STKQELIAALEKGSFLPDYPQKODQFVAEYEPVLIEILVEWMDPSFVCLKIGACPSAH 485  
QY 481 KPLGTEKCIWGPSYWCNTETAACNAVEHCKRHVN 518  
Db 486 KPLGTEKCIWGPSYWCNTETAACNAVEHCKRHVN 523

RESULT 7  
US-09-076-258A-2

; Sequence 2, Application US/09076258A  
; Patent No. 6559124

## GENERAL INFORMATION:

; APPLICANT: O'Brien, John S.

; APPLICANT: Kishimoto, Yasuo

; TITLE OF INVENTION: PHARMACEUTICAL COMPOSITIONS

; TITLE OF INVENTION: COMPRISING PROSAPOSIN AND NEUROTROPHIC PEPTIDES DERIVED

; NUMBER OF SEQUENCES: 11

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Knobbe, Martens, Olson and Bear

; STREET: 620 Newport Center Blvd. 16th Floor

; CITY: Newport Beach

; STATE: CA

; COUNTRY: USA

; ZIP: 92660

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Diskette

; COMPUTER: IBM Compatible

; OPERATING SYSTEM: DOS

; SOFTWARE: FASTSEQ for Windows Version 2.0

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/09/076, 258A

; FILING DATE:

; CLASSIFICATION:

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: 08/958, 970

; FILING DATE: 28-OCT-97

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: 08/483, 146

; FILING DATE: 07-JUN-1995

; APPLICATION NUMBER: 08/100, 247

; FILING DATE: 30-JUL-1993

; ATTORNEY/AGENT INFORMATION:

; NAME: Israelien, Ned A

; REGISTRATION NUMBER: 29, 655

; REFERENCE/DOCKET NUMBER: MYELOS.2DVIC2

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: 619-235-8550

; TELEFAX: 619-235-0176

; TELETYPE:

; INFORMATION FOR SEQ ID NO: 2:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 523 amino acids

; TYPE: amino acid

; STRANDEDNESS: single

; TOPOLOGY: linear

; MOLECULE TYPE: peptide

; FRAGMENT TYPE: N-terminal

## US-09-076-258A-2

## Query Match

Best Local Similarity 98.9%; Score 2727.5; DB 4; Length 523;

Matches 514; Conservative 2; Mismatches 1; Indels 1; Gaps 1;

QY 1 LASLIGALAGPVLGKECTRGSAVWCNVYASDCGAVKCLQTVNKPVTYSLPCIC 60  
Db 7 LASLIGALAGPVLGKECTRGSAVWCNVYASDCGAVKCLQTVNKPVTYSLPCIC 66  
QY 61 KDVTYAGDMKDNATEBEILVYLEKTCMDLPKRNMSASCKEIVDSYLPVLLIIKGMS 120  
Db 67 KDVTYAGDMKDNATEBEILVYLEKTCMDLPKRNMSASCKEIVDSYLPVLLIIKGMS 126  
QY 121 RPGEVCSALNLCESLQKHLAELNHQKLESKIKIIBELDTEVYAAPPMANIPILLYPQDPR 180  
Db 127 RPGEVCSALNLCESLQKHLAELNHQKLESKIKIIBELDTEVYAAPPMANIPILLYPQDPR 186  
QY 181 SKPPKNGDVCDQCIQWVTDIQTAVRTNSTFVQALVEHVKEBCDRLGPGMADICXNYIS 240  
Db 187 SKPPKNGDVCDQCIQWVTDIQTAVRTNSTFVQALVEHVKEBCDRLGPGMADICXNYIS 245  
QY 241 QYSEIAIQMMHMQKEICALVGFCDVEKEMPMQTLVPAKVASKNVLPALVEPIKKE 300  
Db 246 QYSEIAIQMMHMQKEICALVGFCDVEKEMPMQTLVPAKVASKNVLPALVEPIKKE 305  
QY 301 VPAKSDVYCEVCEFLVKEVTYKLIINNKTKEKILDAFDMKCSKLPKSLSEECQEVVDITYGS 360  
Db 306 VPAKSDVYCEVCEFLVKEVTYKLIINNKTKEKILDAFDMKCSKLPKSLSEECQEVVDITYGS 365  
QY 361 SILSLILEEVSPELVCSMLHLCSTGRLPALVHTVHTOPKDGFCVCKLVGYLDRLNEKN 420  
Db 366 SILSLILEEVSPELVCSMLHLCSTGRLPALVHTVHTOPKDGFCVCKLVGYLDRLNEKN 425  
QY 421 STKQELIAALEKGSFLPDYPQKODQFVAEYEPVLIEILVEWMDPSFVCLKIGACPSAH 480  
Db 426 STKQELIAALEKGSFLPDYPQKODQFVAEYEPVLIEILVEWMDPSFVCLKIGACPSAH 485  
QY 481 KPLGTEKCIWGPSYWCNTETAACNAVEHCKRHVN 518  
Db 486 KPLGTEKCIWGPSYWCNTETAACNAVEHCKRHVN 523

## RESULT 8

US-09-352-548-2  
; Sequence 2, Application US/09352548  
; Patent No. 6500431

## GENERAL INFORMATION:

; APPLICANT: Gill, Parkash S.

; APPLICANT: Parkash S. Gill, M.D., Inc.

; TITLE OF INVENTION: No. 6500431el Inhibitors of Angiogenesis and Tumor Growth

; FILE REFERENCE: 017986-000410US

; CURRENT APPLICATION NUMBER: US/09/352, 548

; FILING DATE: 1999-07-12

; EARLIER APPLICATION NUMBER: US 60/092, 647

; FILING DATE: 1998-07-13

; NUMBER OF SEQ ID NOS: 59

; SOFTWARE: Patent In Ver. 2.1

; SEQ ID NO 2

; LENGTH: 81

; TYPE: PRT

; ORGANISM: Homo sapiens

; FEATURE:

; OTHER INFORMATION: Saposin B

US-09-352-548-2

Query Match 15.6%; Score 431; DB 4; Length 81;

Best Local Similarity 100.0%; Pred. No. 1.7e-31;

Matches 81; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 189 GDVCDQCIQWVTDIQTAVRTNSTFVQALVEHVKEBCDRLGPGMADICXNYISQYSEIAIQ 248  
Db 1 GDVCDQCIQWVTDIQTAVRTNSTFVQALVEHVKEBCDRLGPGMADICXNYISQYSEIAIQ 60

QY 249 MWMHQPKEICATVGFCDVYK 269  
Db 61 MWMHQPKEICATVGFCDVYK 81

## RESULT 9

US-08-584-671-15  
Sequence 15, Application US/08584671  
Patent No. 5910568  
GENERAL INFORMATION:  
APPLICANT: HAMMERSTEDT, ROY H, BARBATO, GUY F,  
TITLE OF INVENTION: MOLECULE INVOLVED IN BINDING OF SPERM  
TITLE OF INVENTION: TO EGG SURFACES AND PROCEDURES FOR USE OF THIS MOLECULE  
NUMBER OF SEQUENCES: 16  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: INTELLECTUAL PROPERTY OFFICE, THE PENNSYLVANIA  
ADDRESSER: STATE UNIVERSITY  
STREET: 113 TECHNOLOGY CENTER  
CITY: UNIVERSITY PARK  
STATE: PENNSYLVANIA  
COUNTRY: UNITED STATES OF AMERICA  
ZIP: 16802-7000  
COMPUTER READABLE FORM:  
MEDIUM TYPE: FLOPPY DISK  
COMPUTER: NEC 286  
OPERATING SYSTEM: DOS  
SOFTWARE: WORDPERFECT 5.1  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/584,671  
FILING DATE:  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: MONAHAN, THOMAS J  
REGISTRATION NUMBER: 29835  
REFERENCE/DOCKET NUMBER:  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 814-865-6277  
TELEFAX: 814-865-3591  
INFORMATION FOR SEQ ID NO: 15:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 80  
TYPE: AMINO ACID  
STRANDEDNESS: SINGLE  
TOPOLOGY: UNKNOWN  
US-08-584-671-15

Query Match 15.2%; Score 418.5; DB 2; Length 80;  
Best Local Similarity 98.8%; Pred. No. 2.3e-30;

Matches 80; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

QY 122 PGEVCSALNLCESLQKHLAEINHOQLESNKIPELDMTEVAVPFWANIPLLLYPDGPRS 181  
Db 1 PGEVC-ALNLCESLQKHLAEINHOQLESNKIPELDMTEVAVPFWANIPLLLYPDGPRS 59  
QY 182 KPQPKNDGVCDGCIQWVTDI 202  
Db 60 KPQPKNDGVCDGCIQWVTDI 80

## RESULT 10

US-09-027-376-15  
Sequence 15, Application US/09027376  
Patent No. 6004586  
GENERAL INFORMATION:  
APPLICANT: HAMMERSTEDT, ROY H, BARBATO, GUY F,  
TITLE OF INVENTION: MOLECULE INVOLVED IN BINDING OF SPERM  
TITLE OF INVENTION: TO EGG SURFACES AND PROCEDURES FOR USE OF THIS MOLECULE  
NUMBER OF SEQUENCES: 16

## CORRESPONDENCE ADDRESS:

ADDRESSEE: INTELLECTUAL PROPERTY OFFICE, THE PENNSYLVANIA  
ADDRESSER: STATE UNIVERSITY  
STREET: 113 TECHNOLOGY CENTER  
CITY: UNIVERSITY PARK  
STATE: PENNSYLVANIA  
COUNTRY: UNITED STATES OF AMERICA  
ZIP: 16802-7000  
COMPUTER READABLE FORM:  
MEDIUM TYPE: FLOPPY DISK  
COMPUTER: NEC 286  
OPERATING SYSTEM: DOS  
SOFTWARE: WORDPERFECT 5.1  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/027,376  
FILING DATE:  
CLASSIFICATION:  
Prior Application Data:  
APPLICATION NUMBER: 08/584,671  
FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: MONAHAN, THOMAS J  
REGISTRATION NUMBER: 29835  
REFERENCE/DOCKET NUMBER:  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 814-865-6277  
TELEFAX: 814-865-3591  
INFORMATION FOR SEQ ID NO: 15:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 80  
TYPE: AMINO ACID  
STRANDEDNESS: SINGLE  
TOPOLOGY: UNKNOWN  
US-09-027-376-15

Query Match 15.2%; Score 418.5; DB 3; Length 80;  
Best Local Similarity 98.8%; Pred. No. 2.3e-30;  
Matches 80; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

QY 122 PGEVCSALNLCESLQKHLAEINHOQLESNKIPELDMTEVAVPFWANIPLLLYPDGPRS 181  
Db 1 PGEVC-ALNLCESLQKHLAEINHOQLESNKIPELDMTEVAVPFWANIPLLLYPDGPRS 59  
QY 182 KPQPKNDGVCDGCIQWVTDI 202  
Db 60 KPQPKNDGVCDGCIQWVTDI 80

## RESULT 11

US-09-094-192-15  
Sequence 15, Application US/09094192  
Patent No. 6103483  
GENERAL INFORMATION:  
APPLICANT: HAMMERSTEDT, ROY H, BARBATO, GUY F,  
TITLE OF INVENTION: MOLECULE INVOLVED IN BINDING OF SPERM TO EGG SURFACES AND PROCI  
NUMBER OF SEQUENCES: 16  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: INTELLECTUAL PROPERTY OFFICE, THE PENNSYLVANIA STATE UNIVERSITY  
ADDRESSER: 113 TECHNOLOGY CENTER  
CITY: UNIVERSITY PARK  
STATE: PENNSYLVANIA  
COUNTRY: UNITED STATES OF AMERICA  
ZIP: 16802-7000  
COMPUTER READABLE FORM:  
MEDIUM TYPE: FLOPPY DISK  
COMPUTER: NEC 286  
OPERATING SYSTEM: DOS  
SOFTWARE: WORDPERFECT 5.1  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/094,192  
FILING DATE:  
CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:  
NAME: MONAHAN, THOMAS J  
REGISTRATION NUMBER: 29835  
REFERENCE/DOCKET NUMBER:  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 814-865-6277  
TELEFAX: 814-865-3591  
INFORMATION FOR SEQ ID NO: 15:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 80  
TYPE: AMINO ACID  
STRANDEDNESS: SINGLE  
TOPOLOGY: UNKNOWN  
US-09-094-192-15

Query Match 15.1%; Score 415.5; DB 3; Length 80;  
Best Local Similarity 97.5%; Pred. No. 4.3e-30;  
Matches 79; Conservative 1; Mismatches 0; Indels 1; Gaps 1;

Qy 122 PGEVCSALNLCESLQKHLAEINHOQLSNNKIPELDMTEVVAPEFMANIPLLIYPQDGPFS 181  
Db 1 PGEVCSALNLCESLQKHLAEINHOQLSNNKIPELDMTEVVAPEFMANIPLLIYPQDGPFS 59  
Qy 182 KPQKNGDVCCDCIQWVTDI 202  
Db 60 KPQKNGDVCCDCIQWVTDI 80

RESULT 12  
US-08-100-247-3  
Sequence 3, Application US/08100247  
Patent No. 5571787

GENERAL INFORMATION:

APPLICANT: O'BRIEN, JOHN S.  
APPLICANT: KISHIMOTO, YASUO  
TITLE OF INVENTION: PROSAPOSIN AS A NEUROTROPIC FACTOR  
NUMBER OF SEQUENCES: 5  
CORRESPONDENCE ADDRESS:  
ADDRESS: KNOBBE, MARTENS, OLSON AND BEAR  
STREET: 620 NEWPORT CENTER DRIVE SIXTEENTH FLOOR  
CITY: NEWPORT BEACH  
STATE: CA  
COUNTRY: USA  
ZIP: 92660

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/100.247  
FILING DATE: 19930730

CLASSIFICATION: 514

ATTORNEY/AGENT INFORMATION:

NAME: Israelien, Ned A.  
REGISTRATION NUMBER: 29.655  
REFERENCE/DOCKET NUMBER: O'BRIEN.002A  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 619-235-8550  
TELEFAX: 619-235-0176

INFORMATION FOR SEQ ID NO: 3:

SEQUENCE CHARACTERISTICS:

LENGTH: 80 amino acids

TYPE: amino acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: protein

HYPOTHEICAL: NO

ANTI-SENSE: NO

FRAGMENT TYPE: internal

IMMEDIATE SOURCE:

CLONE: SAPOSIN C  
US-08-100-247-3

Query Match 14.9%; Score 412; DB 1; Length 80;  
Best Local Similarity 100.0%; Pred. No. 8.8e-30;  
Matches 80; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 305 SDVYCEVEFLVKEVTKLIDNNKTEKEILDAFDKXCSLPSLSBECQEVVDYGSSTLS 364  
Db 1 SDVYCEVEFLVKEVTKLIDNNKTEKEILDAFDKXCSLPSLSBECQEVVDYGSSTLS 60

Qy 365 ILLEEVSPELVCSMLHLCSG 384  
Db 61 ILLEEVSPELVCSMLHLCSG 80

RESULT 13  
US-08-483-146A-3

Sequence 3, Application US/08483146A  
Patent No. 5696080

GENERAL INFORMATION:

APPLICANT: O'Brien, John S.  
APPLICANT: Kishimoto, Yasuo  
TITLE OF INVENTION: PHARMACEUTICAL COMPOSITIONS  
TITLE OF INVENTION: COMPRISING PROSAPOSIN AND NEUROTROPIC PEPTIDES DERIVED  
FROM  
NUMBER OF SEQUENCES: 11  
CORRESPONDENCE ADDRESS:  
ADDRESS: Knobbe, Martens, Olson and Bear  
STREET: 620 Newport Center Blvd. 16th Floor  
CITY: Newport Beach  
STATE: CA  
COUNTRY: USA  
ZIP: 92660

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: DOS  
SOFTWARE: FastSeq for Windows Version 2.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/483,146A  
FILING DATE: 07-JUN-1995  
CLASSIFICATION: 514  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER:  
FILING DATE:

ATTORNEY/AGENT INFORMATION:

NAME: Israelien, Ned A.  
REGISTRATION NUMBER: 29.655  
REFERENCE/DOCKET NUMBER: MYELOS.002DV1  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 619-235-8550  
TELEFAX: 619-235-0176

INFORMATION FOR SEQ ID NO: 3:

SEQUENCE CHARACTERISTICS:

LENGTH: 80 amino acids

TYPE: amino acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: protein

FRAGMENT TYPE: internal  
US-08-483-146A-3

Query Match 14.9%; Score 412; DB 1; Length 80;  
Best Local Similarity 100.0%; Pred. No. 8.8e-30;  
Matches 80; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 305 SDVYCEVEFLVKEVTKLIDNNKTEKEILDAFDKXCSLPSLSBECQEVVDYGSSTLS 364  
Db 1 SDVYCEVEFLVKEVTKLIDNNKTEKEILDAFDKXCSLPSLSBECQEVVDYGSSTLS 60

Qy 365 ILLEEVSPELVCSMLHLCSG 384  
Db 61 ILLEEVSPELVCSMLHLCSG 80

RESULT 14  
US-08-232-513A-4  
; Sequence 4, Application US/08232513A  
; Patent No. 570909  
; GENERAL INFORMATION:  
; APPLICANT: O'Brien, John S.  
; TITLE OF INVENTION: Prosaposin and Cytokine-Derived Peptides  
; TITLE OF INVENTION: as Therapeutic Agents  
; NUMBER OF SEQUENCES: 20  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Campbell & Flores LLP  
; STREET: 4370 La Jolla Village Drive, Suite 700  
; CITY: San Diego  
; STATE: California  
; COUNTRY: USA  
; ZIP: 92122  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/232,513A  
; FILING DATE: 21-APR-1994  
; CLASSIFICATION: 514  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/100,247  
; FILING DATE: 30-JUL-1993  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Campbell, Cathryn A.  
; REGISTRATION NUMBER: 31,815  
; REFERENCE/DOCKET NUMBER: P-UD 1643  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (619) 535-8949  
; TELEFAX: (619) 535-9001  
; INFORMATION FOR SEQ ID NO: 4:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 80 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
; FEATURE:  
; NAME/KEY: Peptide  
; LOCATION: 1..80  
; OTHER INFORMATION: /label= Saposin\_C  
US-08-232-513A-4  
Query Match 14.9%; Score 412; DB 1; Length 80;  
Best Local Similarity 100.0%; Pred. No. 8.8e-30;  
Matches 80; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 305 SDVYCEVCEFLVKEVTYKLIIDNNKTEKEILDAFDKXCSKLPKSLSECEVVDYSSILS 364  
DB 1 SDVYCEVCEFLVKEVTYKLIIDNNKTEKEILDAFDKXCSKLPKSLSECEVVDYSSILS 60  
QY 365 ILLEVSPELVCSMLHLCSG 384  
DB 61 ILLEVSPELVCSMLHLCSG 80  
RESULT 15  
US-08-484-594A-3  
; Sequence 3, Application US/08484594A  
; Patent No. 5714459  
; GENERAL INFORMATION:  
; APPLICANT: O'Brien, John S.  
; APPLICANT: Kishimoto, Yasuo  
; TITLE OF INVENTION: USE OF PROSAPOSIN AND NEUROTROPHIC PEPTIDES  
; TITLE OF INVENTION: DERIVED THEREFROM  
; NUMBER OF SEQUENCES: 11  
; CORRESPONDENCE ADDRESS:

ADDRESSEE: Knobbe, Martens, Olson and Bear  
STREET: 620 Newport Center Drive, Sixteenth Floor  
CITY: Newport Beach  
STATE: CA  
COUNTRY: USA  
ZIP: 92660  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette  
; COMPUTER: IBM Compatible  
; OPERATING SYSTEM: DOS  
; SOFTWARE: FastSeq for Windows Version 2.0  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/484,594A  
; FILING DATE: 07-JUN-1995  
; CLASSIFICATION: 514  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/100,247  
; FILING DATE: 30-JUL-1993  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Israelien, Ned A.  
; REGISTRATION NUMBER: 29,655  
; REFERENCE/DOCKET NUMBER: MYELOS.002DV2  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 619-235-8550  
; TELEFAX: 619-235-0176  
; TELEX:  
; INFORMATION FOR SEQ ID NO: 3:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 80 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; FRAGMENT TYPE: Internal  
US-08-484-594A-3  
Query Match 14.9%; Score 412; DB 1; Length 80;  
Best Local Similarity 100.0%; Pred. No. 8.8e-30;  
Matches 80; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 305 SDVYCEVCEFLVKEVTYKLIIDNNKTEKEILDAFDKXCSKLPKSLSECEVVDYSSILS 364  
DB 1 SDVYCEVCEFLVKEVTYKLIIDNNKTEKEILDAFDKXCSKLPKSLSECEVVDYSSILS 60  
QY 365 ILLEVSPELVCSMLHLCSG 384  
DB 61 ILLEVSPELVCSMLHLCSG 80

Search completed: May 5, 2004, 13:16:44  
Job time : 23.6287 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

## OM protein - protein search, using sw model

Run on: May 5, 2004, 13:14:53 ; Search time 55.0959 Seconds  
(without alignments)  
2606.077 Million cell updates/sec

Title: US-09-743-684a-1\_COPY\_7\_524

Perfect score: 2759  
Sequence: 1 LASLIGALAPVGLKECT.....NTETAAQCNAYEHCHRYWN 518

Scoring table: BLOSUM62

Gapop 10.0, Gapext 0.5

Searched: 1138120 seqs, 27189581 residues

Total number of hits satisfying chosen parameters: 1138120

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Database : Published Applications AA:\*

1:	/cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep:*
2:	/cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pep:*
3:	/cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep:*
4:	/cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep:*
5:	/cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep:*
6:	/cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB.pep:*
7:	/cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep:*
8:	/cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep:*
9:	/cgn2_6/ptodata/2/pubpaa/US09A_PUBCOMB.pep:*
10:	/cgn2_6/ptodata/2/pubpaa/US09B_PUBCOMB.pep:*
11:	/cgn2_6/ptodata/2/pubpaa/US09C_PUBCOMB.pep:*
12:	/cgn2_6/ptodata/2/pubpaa/US09C_NEW_PUB.pep:*
13:	/cgn2_6/ptodata/2/pubpaa/US10A_PUBCOMB.pep:*
14:	/cgn2_6/ptodata/2/pubpaa/US10B_PUBCOMB.pep:*
15:	/cgn2_6/ptodata/2/pubpaa/US10C_PUBCOMB.pep:*
16:	/cgn2_6/ptodata/2/pubpaa/US10C_NEW_PUB.pep:*
17:	/cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep:*
18:	/cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2759	100.0	524	9	US-09-870-759-60
2	2759	100.0	524	10	US-09-751-708A-60
3	2759	100.0	524	12	US-10-267-502-386
4	2747.5	99.6	527	9	US-09-870-759-61
5	2747.5	99.6	527	10	US-09-751-708A-61
6	2747.5	99.6	527	14	US-10-060-036-73
7	2727.5	98.9	523	9	US-09-767-007A-2
8	2419.5	87.7	479	10	US-09-978-418A-40
9	1947	70.6	554	14	US-10-205-194-176
10	1189	43.1	521	15	US-10-276-162-1
11	1127.5	40.9	531	11	US-09-833-245-303
12	1093	38.6	209	14	US-10-043-467-340
13	822.5	29.8	362	12	US-10-332-426-8
14	549.5	19.9	953	12	US-10-267-502-385
15	545	19.8	241	15	US-10-108-260A-4529

16	539.5	19.6	156	12	US-09-925-298-644	Sequence 644, App
17	539.5	19.6	156	14	US-10-102-806-644	Sequence 644, App
18	412	14.9	80	9	US-09-767-007A-3	Sequence 3, Appl1
19	412	14.9	80	9	US-09-753-126-3	Sequence 3, Appl1
20	412	14.9	80	15	US-10-330-697-3	Sequence 3, Appl1
21	412	14.9	592	9	US-09-753-126-4	Sequence 4, Appl1
22	412	14.9	592	15	US-10-330-697-4	Sequence 4, Appl1
23	340.5	12.3	381	15	US-10-236-031B-10	Sequence 10, Appl1
24	340.5	12.3	381	15	US-10-295-027-1223	Sequence 1223, Ap
25	235.5	8.3	216	12	US-10-424-599-276328	Sequence 276328, Ap
26	211	7.6	40	10	US-09-780-438A-1	Sequence 1, Appl1
27	211	7.6	212	12	US-10-424-599-157904	Sequence 157904, A
28	204.5	7.4	246	12	US-10-425-114-41077	Sequence 41077, A
29	201	7.3	38	10	US-09-780-438A-2	Sequence 2, Appl1
30	201	7.3	212	12	US-10-424-599-157906	Sequence 157906, A
31	197	7.1	252	12	US-10-425-114-55606	Sequence 55606, A
32	184	6.7	243	12	US-10-425-114-67929	Sequence 67929, A
33	175.5	6.4	265	12	US-10-425-114-51535	Sequence 51535, A
34	169.5	6.1	236	12	US-10-424-599-227975	Sequence 227975, A
35	168.5	6.1	181	8	US-08-488-123-12	Sequence 12, Appl1
36	165.5	6.0	246	12	US-10-425-114-60351	Sequence 60351, A
37	163	5.9	514	14	US-10-339-351-3	Sequence 3, Appl1
38	162	5.9	514	14	US-10-339-351-1	Sequence 1, Appl1
39	159.5	5.8	281	12	US-10-424-599-205602	Sequence 205602, A
40	157.5	5.7	282	12	US-10-424-599-205603	Sequence 205603, A
41	156.5	5.7	507	12	US-10-424-599-245585	Sequence 245585, A
42	152	5.5	471	12	US-10-425-114-71427	Sequence 71427, A
43	151	5.5	393	12	US-10-425-114-65142	Sequence 65142, A
44	147.5	5.3	223	12	US-10-425-114-40454	Sequence 40454, A
45	146.5	5.3	295	12	US-10-425-114-57134	Sequence 57134, A

## ALIGNMENTS

RESULT 1									
US-09-870-759-60									
Sequence 60, Application US/09870759									
Patent No. US20020177551A1									
GENERAL INFORMATION:									
APPLICANT: TERMAN, David S									
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATMENT OF NEOPLASTIC DISEASE									
FILE REFERENCE: 870759									
CURRENT FILING DATE: 2002-01-14									
PRIOR APPLICATION NUMBER: US 09/870,759									
PRIOR FILING DATE: 2000-05-30									
NUMBER OF SEQ ID NOS: 166									
SOFTWARE: PatentIn version 3.1									
SEQ ID NO 60									
LENGTH: 524									
TYPE: PRT									
ORGANISM: Homo sapiens									
US-09-870-759-60									
Query Match									
Best Local Similarity 100.0%; Score 2759; DB 9; Length 524;									
Matches 518; Conservative 0; Mismatches 0; Indels 0; Gaps 0;									
QY	1	LASLIGALAPVGLKECTGSAVWCONVTASDCGAKHCLQTVNKKPTVKSIPCCIC	60						
DB	7	LASLIGALAPVGLKECTGSAVWCONVTASDCGAKHCLQTVNKKPTVKSIPCCIC	66						
QY	61	KDVVTAAGDMKDNATEEEIIVLEKTCMDLKPKNMSASCKEIVDSYLPVILDIKSGMS	120						
DB	67	KDVVTAAGDMKDNATEEEIIVLEKTCMDLKPKNMSASCKEIVDSYLPVILDIKSGMS	126						
QY	121	RPEVCSTLNLCESIQKHLAINHOKLESKIKIPELMTETVAPPMANIPILLVPQDSR	180						
DB	127	RPEVCSTLNLCESIQKHLAINHOKLESKIKIPELMTETVAPPMANIPILLVPQDSR	186						
QY	181	SKPOPKDGVCDQCICQWTDIQTAVRTNSTFVALVHVHVEBCDRLGPMADICKNYS	240						

```

Db 187 SKPQKNDGVCQDCIQWVTDIQTAVRTNSTFVQALVEHVEECBDRLGPGMADICKNYIS 246
QY 241 QYSEIATQMMHMQPKKEICALVGFCDVEKEMPMQTLVPARVASKNVIPALELVEPIKKEH 300
Db 247 QYSEIATQMMHMQPKKEICALVGFCDVEKEMPMQTLVPARVASKNVIPALELVEPIKKEH 306
QY 301 VPAKSDVYCEVCEFLVEVTKLIDNNKTEKEIILDAFPMKSKLPKSLSEECQEVDTYGS 360
Db 307 VPAKSDVYCEVCEFLVEVTKLIDNNKTEKEIILDAFPMKSKLPKSLSEECQEVDTYGS 366
QY 361 SILSILLEEVSPELVCSMLHCSGTRLPALTVAHTQPKDGFCEVCKLGVYLDRLNEKN 420
Db 367 SILSILLEEVSPELVCSMLHCSGTRLPALTVAHTQPKDGFCEVCKLGVYLDRLNEKN 426
QY 421 STKOEIILAALKEKGSFLPDPYQKQCDQFVAEYEPVLIELIVEWMDSPFVLKIGACPSAH 480
Db 427 STKOEIILAALKEKGSFLPDPYQKQCDQFVAEYEPVLIELIVEWMDSPFVLKIGACPSAH 486
QY 481 KPLIGTEKCIWGPSYWCQNTETAACQNAVEHCKRHVWN 518
Db 487 KPLIGTEKCIWGPSYWCQNTETAACQNAVEHCKRHVWN 524

```

## RESULT 2

```

US-09-751-708A-60
; Sequence 60, Application US/09751708A
; Publication No. US20030157113A1
; GENERAL INFORMATION:
; APPLICANT: TERMAN, David S
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATMENT OF NEOPLASTIC DISEASE
; FILE REFERENCE: 751708
; CURRENT APPLICATION NUMBER: US/09/751,708A
; PRIOR FILING DATE: 2002-10-15
; PRIOR APPLICATION NUMBER: US 60/173,371
; PRIOR FILING DATE: 1999-12-28
; NUMBER OF SEQ ID NOS: 166
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 60
; LENGTH: 524
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-751-708A-60

```

```

Query Match      100.0%; Score 2759; DB 10; Length 524;
Best Local Similarity 100.0%; Pred. No. 5,1e-237;
Matches 518; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 1 LASILGAALAGPVGLKECTRGSAVWCQNVKTASDCGAVKHCLQTVWNNKPTVKSILPCDIC 60
Db 7 LASILGAALAGPVGLKECTRGSAVWCQNVKTASDCGAVKHCLQTVWNNKPTVKSILPCDIC 66
QY 61 KDVTAAAGDMLKDNATEEELIIVLEKTCDWLPEKPNMSASCKEIVDSYLPVILDIKIGEMS 120
Db 67 KDVTAAAGDMLKDNATEEELIIVLEKTCDWLPEKPNMSASCKEIVDSYLPVILDIKIGEMS 126
QY 121 RPEVCSALNLCESLQKHLAEINHOKELESNKIPBLDMTEVYVAPPMANIPILLIYPQDGR 180
Db 127 RPEVCSALNLCESLQKHLAEINHOKELESNKIPBLDMTEVYVAPPMANIPILLIYPQDGR 186
QY 181 SKPQKNDGVCQDCIQWVTDIQTAVRTNSTFVQALVEHVEECBDRLGPGMADICKNYIS 240
Db 187 SKPQKNDGVCQDCIQWVTDIQTAVRTNSTFVQALVEHVEECBDRLGPGMADICKNYIS 246
QY 241 QYSEIATQMMHMQPKKEICALVGFCDVEKEMPMQTLVPARVASKNVIPALELVEPIKKEH 300
Db 247 QYSEIATQMMHMQPKKEICALVGFCDVEKEMPMQTLVPARVASKNVIPALELVEPIKKEH 306
QY 301 VPAKSDVYCEVCEFLVEVTKLIDNNKTEKEIILDAFPMKSKLPKSLSEECQEVDTYGS 360
Db 307 VPAKSDVYCEVCEFLVEVTKLIDNNKTEKEIILDAFPMKSKLPKSLSEECQEVDTYGS 366
QY 361 SILSILLEEVSPELVCSMLHCSGTRLPALTVAHTQPKDGFCEVCKLGVYLDRLNEKN 420

```

```

Db 367 SILSILLEEVSPELVCSMLHCSGTRLPALTVAHTQPKDGFCEVCKLGVYLDRLNEKN 426
QY 421 STKOEIILAALKEKGSFLPDPYQKQCDQFVAEYEPVLIELIVEWMDSPFVLKIGACPSAH 480
Db 427 STKOEIILAALKEKGSFLPDPYQKQCDQFVAEYEPVLIELIVEWMDSPFVLKIGACPSAH 486
QY 481 KPLIGTEKCIWGPSYWCQNTETAACQNAVEHCKRHVWN 518
Db 487 KPLIGTEKCIWGPSYWCQNTETAACQNAVEHCKRHVWN 524

```

## RESULT 3

```

US-10-267-502-386
; Sequence 386, Application US/10267502
; Publication No. US20040071700A1
; GENERAL INFORMATION:
; APPLICANT: Kim, Jaeseob
; TITLE OF INVENTION: Obesity Linked Genes
; FILE REFERENCE: LSD-07416
; CURRENT APPLICATION NUMBER: US/10/267,502
; CURRENT FILING DATE: 2003-01-27
; NUMBER OF SEQ ID NOS: 439
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 386
; LENGTH: 524
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-267-502-386

```

```

Query Match      100.0%; Score 2759; DB 12; Length 524;
Best Local Similarity 100.0%; Pred. No. 5,1e-237;
Matches 518; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 1 LASILGAALAGPVGLKECTRGSAVWCQNVKTASDCGAVKHCLQTVWNNKPTVKSILPCDIC 60
Db 7 LASILGAALAGPVGLKECTRGSAVWCQNVKTASDCGAVKHCLQTVWNNKPTVKSILPCDIC 66
QY 61 KDVTAAAGDMLKDNATEEELIIVLEKTCDWLPEKPNMSASCKEIVDSYLPVILDIKIGEMS 120
Db 67 KDVTAAAGDMLKDNATEEELIIVLEKTCDWLPEKPNMSASCKEIVDSYLPVILDIKIGEMS 126
QY 121 RPEVCSALNLCESLQKHLAEINHOKELESNKIPBLDMTEVYVAPPMANIPILLIYPQDGR 180
Db 127 RPEVCSALNLCESLQKHLAEINHOKELESNKIPBLDMTEVYVAPPMANIPILLIYPQDGR 186
QY 181 SKPQKNDGVCQDCIQWVTDIQTAVRTNSTFVQALVEHVEECBDRLGPGMADICKNYIS 240
Db 187 SKPQKNDGVCQDCIQWVTDIQTAVRTNSTFVQALVEHVEECBDRLGPGMADICKNYIS 246
QY 241 QYSEIATQMMHMQPKKEICALVGFCDVEKEMPMQTLVPARVASKNVIPALELVEPIKKEH 300
Db 247 QYSEIATQMMHMQPKKEICALVGFCDVEKEMPMQTLVPARVASKNVIPALELVEPIKKEH 306
QY 301 VPAKSDVYCEVCEFLVEVTKLIDNNKTEKEIILDAFPMKSKLPKSLSEECQEVDTYGS 360
Db 307 VPAKSDVYCEVCEFLVEVTKLIDNNKTEKEIILDAFPMKSKLPKSLSEECQEVDTYGS 366
QY 361 SILSILLEEVSPELVCSMLHCSGTRLPALTVAHTQPKDGFCEVCKLGVYLDRLNEKN 420
Db 367 SILSILLEEVSPELVCSMLHCSGTRLPALTVAHTQPKDGFCEVCKLGVYLDRLNEKN 426
QY 421 STKOEIILAALKEKGSFLPDPYQKQCDQFVAEYEPVLIELIVEWMDSPFVLKIGACPSAH 480
Db 427 STKOEIILAALKEKGSFLPDPYQKQCDQFVAEYEPVLIELIVEWMDSPFVLKIGACPSAH 486
QY 481 KPLIGTEKCIWGPSYWCQNTETAACQNAVEHCKRHVWN 518
Db 487 KPLIGTEKCIWGPSYWCQNTETAACQNAVEHCKRHVWN 524

```

## RESULT 4

```
US-09-870-759-61
```



```
/ Sequence 61, Application US/09870759
/ Patent No. US20020177551A1
/ GENERAL INFORMATION:
/ APPLICANT: TERMAN, David S
/ TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATMENT OF NEOPLASTIC DISEASE
/ FILE REFERENCE: 870759
/ CURRENT APPLICATION NUMBER: US/09/870,759
/ CURRENT FILING DATE: 2002-01-14
/ PRIOR APPLICATION NUMBER: US 60/208,128
/ PRIOR FILING DATE: 2000-05-30
/ NUMBER OF SEQ ID NOS: 166
/ SOFTWARE: PatentIn version 3.1
/ SEQ ID NO 61
/ LENGTH: 527
/ TYPE: PRT
/ ORGANISM: Homo sapiens
US-09-870-759-61

Query Match          99.6%; Score 2747.5; DB 9; Length 527;
Best Local Similarity 99.4%; Pred. No. 5.4e-236;
Matches 518; Conservative 0; Mismatches 0; Indels 3; Gaps 1;

QY 1 LASLIGALAGPVIGLKECTRGSAVWCQNVKTASDCGAVKHCLQTVWKKPTVKSIPCDIC 60
DB 7 LASLIGALAGPVIGLKECTRGSAVWCQNVKTASDCGAVKHCLQTVWKKPTVKSIPCDIC 66
QY 61 KDVTAAAGDMLKDNTAEELIIVLEKTCDDMLPKPMNSASCKEIVDSYLPVLLDIIKGMS 120
DB 67 KDVTAAAGDMLKDNTAEELIIVLEKTCDDMLPKPMNSASCKEIVDSYLPVLLDIIKGMS 126
QY 121 RPEVCSALNLCESLQKHLAEINHOQKLESNKIPELDMTEVVAAPPMANIPLLIYPQDGR 180
DB 127 RPEVCSALNLCESLQKHLAEINHOQKLESNKIPELDMTEVVAAPPMANIPLLIYPQDGR 186
QY 181 SKPQKNDGVCQDCIQWVTDIQTAVRTNSTFVQALVEHVKECDRLGPGMADICKNYIS 240
DB 187 SKPQKNDGVCQDCIQWVTDIQTAVRTNSTFVQALVEHVKECDRLGPGMADICKNYIS 246
QY 241 QYSEIAIOMMMHMM--OPKEICALVGFCDVEKEMPMQTLVPAKVASKNVIALLEVEPIK 297
DB 247 QYSEIAIOMMMHMMQDOQPEKICALVGFCDVEKEMPMQTLVPAKVASKNVIALLEVEPIK 306
QY 298 KHEVPAKSDVYCEVCEFLVKEVTKLIDNNKTEKEILDAFDKMSCKLPSLSSECOEVDVT 357
DB 307 KHEVPAKSDVYCEVCEFLVKEVTKLIDNNKTEKEILDAFDKMSCKLPSLSSECOEVDVT 366
QY 358 YGSSILSLILEEVSPELVCSMLHLCSGTRLPALIVHTVTPKXGDFCEVCKLVGYLDNRL 417
DB 367 YGSSILSLILEEVSPELVCSMLHLCSGTRLPALIVHTVTPKXGDFCEVCKLVGYLDNRL 426
QY 418 EKNSTKQEIILALEKGSFLLPDPYQKCDQFAEIEPVLIETLVEWMPSPVCLKIGACP 477
DB 427 EKNSTKQEIILALEKGSFLLPDPYQKCDQFAEIEPVLIETLVEWMPSPVCLKIGACP 486
QY 478 SAHKPLGTEKICMGPSYWCNTETAAQCNAVEHCKRHVNN 518
DB 487 SAHKPLGTEKICMGPSYWCNTETAAQCNAVEHCKRHVNN 527

RESULT 5
US-09-751-708A-61
/ Sequence 61, Application US/09751708A
/ Publication No. US20030157113A1
/ GENERAL INFORMATION:
/ APPLICANT: TERMAN, David S
/ TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATMENT OF NEOPLASTIC DISEASE
/ FILE REFERENCE: 751708
/ CURRENT APPLICATION NUMBER: US/09/751,708A
/ CURRENT FILING DATE: 2002-10-15
/ PRIOR APPLICATION NUMBER: US 60/173,371
/ PRIOR FILING DATE: 1999-12-28
/ NUMBER OF SEQ ID NOS: 166
/ SOFTWARE: PatentIn version 3.1
```

```
/ SEQ ID NO 61
/ LENGTH: 527
/ TYPE: PRT
/ ORGANISM: Homo sapiens
US-09-751-708A-61

Query Match          99.6%; Score 2747.5; DB 10; Length 527;
Best Local Similarity 99.4%; Pred. No. 5.4e-236;
Matches 518; Conservative 0; Mismatches 0; Indels 3; Gaps 1;

QY 1 LASLIGALAGPVIGLKECTRGSAVWCQNVKTASDCGAVKHCLQTVWKKPTVKSIPCDIC 60
DB 7 LASLIGALAGPVIGLKECTRGSAVWCQNVKTASDCGAVKHCLQTVWKKPTVKSIPCDIC 66
QY 61 KDVTAAAGDMLKDNTAEELIIVLEKTCDDMLPKPMNSASCKEIVDSYLPVLLDIIKGMS 120
DB 67 KDVTAAAGDMLKDNTAEELIIVLEKTCDDMLPKPMNSASCKEIVDSYLPVLLDIIKGMS 126
QY 121 RPEVCSALNLCESLQKHLAEINHOQKLESNKIPELDMTEVVAAPPMANIPLLIYPQDGR 180
DB 127 RPEVCSALNLCESLQKHLAEINHOQKLESNKIPELDMTEVVAAPPMANIPLLIYPQDGR 186
QY 181 SKPQKNDGVCQDCIQWVTDIQTAVRTNSTFVQALVEHVKECDRLGPGMADICKNYIS 240
DB 187 SKPQKNDGVCQDCIQWVTDIQTAVRTNSTFVQALVEHVKECDRLGPGMADICKNYIS 246
QY 241 QYSEIAIOMMMHMM--OPKEICALVGFCDVEKEMPMQTLVPAKVASKNVIALLEVEPIK 297
DB 247 QYSEIAIOMMMHMMQDOQPEKICALVGFCDVEKEMPMQTLVPAKVASKNVIALLEVEPIK 306
QY 298 KHEVPAKSDVYCEVCEFLVKEVTKLIDNNKTEKEILDAFDKMSCKLPSLSSECOEVDVT 357
DB 307 KHEVPAKSDVYCEVCEFLVKEVTKLIDNNKTEKEILDAFDKMSCKLPSLSSECOEVDVT 366
QY 358 YGSSILSLILEEVSPELVCSMLHLCSGTRLPALIVHTVTPKXGDFCEVCKLVGYLDNRL 417
DB 367 YGSSILSLILEEVSPELVCSMLHLCSGTRLPALIVHTVTPKXGDFCEVCKLVGYLDNRL 426
QY 418 EKNSTKQEIILALEKGSFLLPDPYQKCDQFAEIEPVLIETLVEWMPSPVCLKIGACP 477
DB 427 EKNSTKQEIILALEKGSFLLPDPYQKCDQFAEIEPVLIETLVEWMPSPVCLKIGACP 486
QY 478 SAHKPLGTEKICMGPSYWCNTETAAQCNAVEHCKRHVNN 518
DB 487 SAHKPLGTEKICMGPSYWCNTETAAQCNAVEHCKRHVNN 527

RESULT 6
US-10-060-036-73
/ Sequence 73, Application US/10060036
/ Publication No. US20030073144A1
/ GENERAL INFORMATION:
/ APPLICANT: Benson, Darin R.
/ APPLICANT: Kalos, Michael D.
/ APPLICANT: Lodes, Michael J.
/ APPLICANT: Persing, David H.
/ APPLICANT: Hepler, William T.
/ APPLICANT: Jiang, Yugu
/ TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
/ FILE REFERENCE: 210121.566
/ CURRENT APPLICATION NUMBER: US/10/060,036
/ CURRENT FILING DATE: 2002-01-30
/ NUMBER OF SEQ ID NOS: 4560
/ SOFTWARE: FastSeq for Windows Version 4.0
/ SEQ ID NO 73
/ LENGTH: 527
/ TYPE: PRT
/ ORGANISM: Homo sapiens
US-10-060-036-73

Query Match          99.6%; Score 2747.5; DB 14; Length 527;
Best Local Similarity 99.4%; Pred. No. 5.4e-236;
```

Matches 518; Conservative 0; Mismatches 0; Indels 3; Gaps 1;

```
QY 1 LASLIGALAGPVIGLKECTRGSAVWCQNYKTASDCGAVHGCLOTWNNKFTVSLPCDIC 60
Db 7 LASLIGALAGPVIGLKECTRGSAVWCQNYKTASDCGAVHGCLOTWNNKFTVSLPCDIC 66
QY 61 KDVTAAAGDMLKDNATEEELIVLEKTCDWLPKPNMSASCKEIVDSYLPVILDIIGEMS 120
Db 67 KDVTAAAGDMLKDNATEEELIVLEKTCDWLPKPNMSASCKEIVDSYLPVILDIIGEMS 126
QY 121 RRGVCSALNLCESLQGHIAELNHQKQLESNKIPELDMTEVVAPEMANIPLLIYPDGGPR 180
Db 127 RRGVCSALNLCESLQGHIAELNHQKQLESNKIPELDMTEVVAPEMANIPLLIYPDGGPR 186
QY 181 SKQPKNDGVCCDCIOMVTDIOTAVRTNSTFVQALVEHKECDRLGPMADICKNYIS 240
Db 187 SKQPKNDGVCCDCIOMVTDIOTAVRTNSTFVQALVEHKECDRLGPMADICKNYIS 246
QY 241 QYSEIAIQMMHMM--OPKEICALVGFCEVKEMPQTLVPAVASKNYIPALEVEPIK 297
Db 247 QYSEIAIQMMHMMQDQPKKEICALVGFCEVKEMPQTLVPAVASKNYIPALEVEPIK 306
QY 298 KHEVPAKSDVYCEVCEFLVEVTKLIDNNKTEKEILDADPKMSCKLPKSLSECEGVVDY 357
Db 307 KHEVPAKSDVYCEVCEFLVEVTKLIDNNKTEKEILDADPKMSCKLPKSLSECEGVVDY 366
QY 358 YGSSILSILLEEVSPELVCSMLHLCSTRLPALTVHTOPKGGGFEVCKLVGLDNL 417
Db 367 YGSSILSILLEEVSPELVCSMLHLCSTRLPALTVHTOPKGGGFEVCKLVGLDNL 426
QY 418 EKSTQOEILALEKGCFLPDPYQKCDQFVAEYEPVILEIVEMDPSFVCLKIGACP 477
Db 427 EKSTQOEILALEKGCFLPDPYQKCDQFVAEYEPVILEIVEMDPSFVCLKIGACP 486
QY 478 SAHKPLGTEKCIWGPSYWCNTETAAQCNAVEHCKRHVN 518
Db 487 SAHKPLGTEKCIWGPSYWCNTETAAQCNAVEHCKRHVN 527

RESULT 7
US-09-767-007A-2
; Sequence 2, Application US/09767007A
; Patent No. US20020077275A1
; GENERAL INFORMATION:
; APPLICANT: John S. O'Brien
; APPLICANT: Yasuo Kishimoto
; TITLE OF INVENTION: PHARMACEUTICAL COMPOSITIONS COMPRISING
; FILE REFERENCE: MYELOS-2DC1C1
; CURRENT APPLICATION NUMBER: US/09/767,007A
; PRIOR FILING DATE: 2001-01-22
; PRIOR APPLICATION NUMBER: 08/958,970
; PRIOR FILING DATE: 1997-10-28
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 523
; TYPE: PRT
; ORGANISM: homo sapiens
US-09-767-007A-2
```

Query Match 98.9%; Score 2727.5; DB 9; Length 523;  
Best Local Similarity 99.2%; Pred. No. 3.3e-24;  
Matches 514; Conservative 2; Mismatches 1; Indels 1; Gaps 1;

```
QY 1 LASLIGALAGPVIGLKECTRGSAVWCQNYKTASDCGAVHGCLOTWNNKFTVSLPCDIC 60
Db 7 LASLIGALAGPVIGLKECTRGSAVWCQNYKTASDCGAVHGCLOTWNNKFTVSLPCDIC 66
QY 61 KDVTAAAGDMLKDNATEEELIVLEKTCDWLPKPNMSASCKEIVDSYLPVILDIIGEMS 120
Db 67 KDVTAAAGDMLKDNATEEELIVLEKTCDWLPKPNMSASCKEIVDSYLPVILDIIGEMS 126
```

```
QY 121 RRGVCSALNLCESLQGHIAELNHQKQLESNKIPELDMTEVVAPEMANIPLLIYPDGGPR 180
Db 127 RRGVCSALNLCESLQGHIAELNHQKQLESNKIPELDMTEVVAPEMANIPLLIYPDGGPR 186
QY 181 SKQPKNDGVCCDCIOMVTDIOTAVRTNSTFVQALVEHKECDRLGPMADICKNYIS 240
Db 187 SKQPKNDGVCCDCIOMVTDIOTAVRTNSTFVQALVEHKECDRLGPMADICKNYIS 246
QY 241 QYSEIAIQMMHMMQPKKEICALVGFCEVKEMPQTLVPAVASKNYIPALEVEPIK 300
Db 246 QYSEIAIQMMHMMQPKKEICALVGFCEVKEMPQTLVPAVASKNYIPALEVEPIK 305
QY 301 VPAKSDVYCEVCEFLVEVTKLIDNNKTEKEILDADPKMSCKLPKSLSECEGVVDY 360
Db 306 VPAKSDVYCEVCEFLVEVTKLIDNNKTEKEILDADPKMSCKLPKSLSECEGVVDY 365
QY 361 SLSILLEEVSPELVCSMLHLCSTRLPALTVHTOPKGGGFEVCKLVGLDNL 420
Db 366 SLSILLEEVSPELVCSMLHLCSTRLPALTVHTOPKGGGFEVCKLVGLDNL 425
QY 421 STQOEILALEKGCFLPDPYQKCDQFVAEYEPVILEIVEMDPSFVCLKIGACPSAH 480
Db 426 STQOEILALEKGCFLPDPYQKCDQFVAEYEPVILEIVEMDPSFVCLKIGACPSAH 485
QY 481 KPLGTEKCIWGPSYWCNTETAAQCNAVEHCKRHVN 518
Db 486 KPLGTEKCIWGPSYWCNTETAAQCNAVEHCKRHVN 523
```

```
RESULT 8
US-09-978-418-40
; Sequence 40, Application US/09978418
; Publication No. US20030118997A1
; GENERAL INFORMATION:
; APPLICANT: Benjamin, Stephan
; APPLICANT: Tanaka, Hiroaki
; TITLE OF INVENTION: HUMAN CDNAS AND PROTEINS AND USES THEREOF
; FILE REFERENCE: 142 US5, REG
; CURRENT APPLICATION NUMBER: US/09/978,418
; PRIOR FILING DATE: 2002-10-15
; PRIOR APPLICATION NUMBER: 60/311,305
; PRIOR FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: 60/314,734
; PRIOR FILING DATE: 2001-08-24
; PRIOR APPLICATION NUMBER: 60/318,204
; PRIOR FILING DATE: 2001-09-07
; PRIOR APPLICATION NUMBER: 60/326,470
; PRIOR FILING DATE: 2001-10-01
; NUMBER OF SEQ ID NOS: 52
; SOFTWARE: JPatent
; SEQ ID NO 40
; LENGTH: 479
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-978-418-40
```

Query Match 87.7%; Score 2419.5; DB 10; Length 479;  
Best Local Similarity 99.4%; Pred. No. 8.3e-207;  
Matches 463; Conservative 0; Mismatches 0; Indels 3; Gaps 1;

```
QY 1 LASLIGALAGPVIGLKECTRGSAVWCQNYKTASDCGAVHGCLOTWNNKFTVSLPCDIC 60
Db 7 LASLIGALAGPVIGLKECTRGSAVWCQNYKTASDCGAVHGCLOTWNNKFTVSLPCDIC 66
QY 61 KDVTAAAGDMLKDNATEEELIVLEKTCDWLPKPNMSASCKEIVDSYLPVILDIIGEMS 120
Db 67 KDVTAAAGDMLKDNATEEELIVLEKTCDWLPKPNMSASCKEIVDSYLPVILDIIGEMS 126
QY 121 RRGVCSALNLCESLQGHIAELNHQKQLESNKIPELDMTEVVAPEMANIPLLIYPDGGPR 180
Db 127 RRGVCSALNLCESLQGHIAELNHQKQLESNKIPELDMTEVVAPEMANIPLLIYPDGGPR 186
QY 181 SKQPKNDGVCCDCIOMVTDIOTAVRTNSTFVQALVEHKECDRLGPMADICKNYIS 240
```

Db 187 SKPGKNDGVCQDCIQWTDIQTAVRTNSTFVQALVHEVKEECRCRLPGMADICKNYIS 246  
Qy 241 QYSEIAIOMMMHMM---QPKETCALVGFCDVEYKEMQTLVPAKASKVITPALEIVEIK 297  
Db 247 QYSEIAIOMMMHMMQDOQPEICALVGFCDVEYKEMQTLVPAKASKVITPALEIVEIK 306  
Qy 298 KHEVPAKSDVYCEVCEFLVKEVTKLIDNNKTEKEILDAFDKMSCKLPSLSSECOEVDVT 357  
Db 307 KHEVPAKSDVYCEVCEFLVKEVTKLIDNNKTEKEILDAFDKMSCKLPSLSSECOEVDVT 366  
Qy 358 YGSSILSLIEBVSPELVCSMLHLCSGTRLPALTVHTVQPKDGFCEVCKLVGYLDNRL 417  
Db 367 YGSSILSLIEBVSPELVCSMLHLCSGTRLPALTVHTVQPKDGFCEVCKLVGYLDNRL 426  
Qy 418 EKNSTKOEIILAEKGSFLPPYQKQCDQFVAEPEVLIELVAV 463  
Db 427 EKNSTKOEIILAEKGSFLPPYQKQCDQFVAEPEVLIELVAV 472

## RESULT 9

US-10-205-194-176  
; Sequence 176, Application US/10205194  
; Publication No. US20030134301A1  
; GENERAL INFORMATION:  
; APPLICANT: Warner-Lambert Company  
; APPLICANT: Lee, Kevin  
; APPLICANT: Dixon, Alistair  
; APPLICANT: Brooksbank, Robert  
; APPLICANT: Pimock, Robert  
; TITLE OF INVENTION: Identification and Use of Molecules Implicated in Pain  
; FILE REFERENCE: WL-A-018201  
; CURRENT APPLICATION NUMBER: US/10/205,194  
; PRIOR FILING DATE: 5200-07-24  
; PRIOR APPLICATION NUMBER: GB 0118354.0  
; NUMBER OF SEQ ID NOS: 177  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 176  
; LENGTH: 554  
; TYPE: PRT  
; ORGANISM: Rattus sp.  
; FEATURE:  
; OTHER INFORMATION: Prosaposin  
US-10-205-194-176

Query Match 70.6%; Score 1947; DB 14; Length 554;  
Best Local Similarity 65.6%; Pred. No. 1.4e-164;

Matches 360; Conservative 77; Mismatches 80; Indels 32; Gaps 3;

Qy 1 LASILGALAGPVGLKCTRGSAVWCONVKTASQCAVKGHLQTVVNNKPTVKSIPCTIC 60  
Db 7 LASILVTLVLTSPVQDPKICSGSAAVCDVKTAVDCAVKGQOMVWAKPTAKSLPCDIC 66  
Qy 61 KDVTAAADMLKDNATEEELIVLEKTCMDLPEKPMNSASCKEIVDSYLPVILDIKSGMS 120  
Db 67 KTVVTEAGNLKDNATEEELIHYLEKTCAMIHDSLSASCKEIVDSYLVILDIKSGMS 126  
Qy 121 RPEVCASALNLCESLQKLAELNHOKESNKIPELMDTEVAVPMANIPLLLYPOQDPR 180  
Db 127 NPEVCASALNLCOSLOEYLAEQN-QRLESNKIPREVDLARVAVAPMSNIPLLLYPOQDPR 185  
Qy 181 SKRQPKNDGVCQDCIQWTDIQTAVRTNSTFVQALVHEVKEECRCRLPGMADICKNYIS 240  
Db 186 SQPQPKABEDVQDCMDKVTDTIQTAVRTNSTFVQALVHEVKEECRCRLPGVSDICKNYVD 245  
Qy 241 QYSEIAIOMMMHMMQPEICALVGFCDVEYKEMQTLVPAKASKVITPALEIVEIK 300  
Db 246 QYSEIAIOMMMHMMQPEICALVGFCDVEYKEMQTLVPAKASKVITPALEIVEIK 305  
Qy 301 VPAKSDVYCEVCEFLVKEVTKLIDNNKTEKEILDAFDKMSCKLPSLSSECOEVDVYGS 360  
Db 306 IQAQNVIQCVQCOLVWRKLSLIIINNATEEELIKGLSKACSLPAPASTKQCEVLVTRGP 365

Qy 361 SILSLIEBVSPELVCSMLHLCSG-----TRLP-----A 389  
Db 366 SLILVLEHENVNPNLCVITSLCSANPNLVGTLBPAAAIYSAIPKEPAPKQPEPKQSA 425  
Qy 390 LTVHTVQPKDGFCEVCKLVGYLDNRLKONSTKOEIILAEKGSFLPPYQKQCDQFV 449  
Db 426 LRAHVPPQKNGGFCEVCKLVITYLEHNLKONSTKEEVYLAALKEKGSFLPPYQKQCDQFV 485  
Qy 450 AEVEPVLIELVEMDSFVCLKTGACPSAHKPLLTGTEKCIWGPSYCONMETAAQCNV 509  
Db 486 AEVEPVLIELVEMDSFVCSKIGVCPSAVKLLGTGECVGRGVCYCONMETAAQCNV 545  
Qy 510 EHCRRHYVN 518  
Db 546 DHCRRHYVN 554

## RESULT 10

US-10-276-162-1  
; Sequence 1, Application US/10276162  
; Publication No. US20030215822A1  
; GENERAL INFORMATION:  
; APPLICANT: GRIFFIN, Jennifer A.  
; APPLICANT: YAO, Monique G.  
; APPLICANT: BRUNS, Christopher M.  
; APPLICANT: YUE, Henry  
; APPLICANT: DELEGEANT, Angelo M.  
; APPLICANT: HAFALIA, April  
; APPLICANT: PATTERSON, Chandra  
; APPLICANT: POLICKY, Jennifer L.  
; APPLICANT: TRIBOULEY, Catherine M.  
; APPLICANT: BAUGHN, Mariah R.  
; APPLICANT: NGUYEN, Daniel B.  
; APPLICANT: LAU, Preeti  
; APPLICANT: TANG, Y. Tom  
; APPLICANT: JACKSON, Jennifer L.  
; APPLICANT: LU, Dying Aina M.  
; APPLICANT: BATRA, Sajeev  
; APPLICANT: AU-YOUNG, Janice  
; APPLICANT: REDDY, Roopa  
; APPLICANT: AZIMZAI, Yalda  
; TITLE OF INVENTION: SECRETED PROTEINS  
; FILE REFERENCE: PI-0071 USN  
; CURRENT APPLICATION NUMBER: US/10/276,162  
; PRIOR FILING DATE: 2002-10-15  
; PRIOR APPLICATION NUMBER: US01/11861  
; PRIOR FILING DATE: 2001-04-11  
; PRIOR APPLICATION NUMBER: 60/197,854  
; PRIOR FILING DATE: 2000-04-14  
; PRIOR APPLICATION NUMBER: 60/202,373  
; PRIOR FILING DATE: 2000-05-04  
; PRIOR APPLICATION NUMBER: 60/205,899  
; PRIOR FILING DATE: 2000-05-18  
; PRIOR APPLICATION NUMBER: 60/210,155  
; PRIOR FILING DATE: 2000-06-01  
; PRIOR APPLICATION NUMBER: 60/209,401  
; NUMBER OF SEQ ID NOS: 28  
; SOFTWARE: PERL Program  
; SEQ ID NO 1  
; LENGTH: 521  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: misc feature  
; OTHER INFORMATION: Incyte ID No. US20030215822A1 7473577CD1  
US-10-276-162-1

Query Match 43.1%; Score 1189; DB 15; Length 521;  
Best Local Similarity 44.1%; Pred. No. 4.9e-97;  
Matches 234; Conservative 98; Mismatches 157; Indels 42; Gaps 10;

```

QY 1 LASLIGALAGPVLGLKECTRGSAVMCONVKTASDCGAKHCLQTYWKNKPTVSLACDTC 60
DB 8 LPSLIGATRASPSGPECAKSTVMQODIQOTARCAVGTCCGAANNKPTASLPCDVC 67
QY 61 KDVTAAAGDMLKDNATEEELIVYEKTCMDLPRKNNASCKEIVDSYLPVILDIIGKEM- 119
DB 68 QDIAAAGNGINPATESDITLALVMKTCFWLPQESSAGCKMMVDAHSSAILSLMARGAD 127
QY 120 SRPGEVSALNLCSLQKHLAEINHOQKLESNKIPELDMTEVAVAPPMANIPILLYPODGP 179
DB 128 SAPAQVCTALSLCEPLQRHLATL-----RPLSKEDTEBAVAPPMANIPILFFHROAP 179
QY 180 RSKPQPKNDGVQDCDIQMTDITQAVRTNSTFVQALVEHVEKBCDRLPGMADICKNYI 239
DB 180 -----EGALCDCCVRQVSRLOEAVRSNLTADL---NIQRCESLPGIHLAVLCNLY 228
QY 240 SQYSEIALQMMHMQPKKEICALVGFCDVEKEMQTLVPAK---VASKNVIPLBELVEPI 296
DB 229 FQFVFPADQALRLPPELCEKRGKGFCEELG-----APARLTQVAMDGVPSELGLPR 281
QY 297 KKHVEPAKSDVYCEVCEFLVKEVTKLIDNNKTEKEILDAFDMKCSKLPKSLSEECQEVND 356
DB 282 KQSEMOMKAGVTCVCECMNVVQKLDHWMNSSELMITHALERCSVMPASITKECIIIVD 341
QY 357 TYGSSILSLLEESPELVCSMLHLCSGTRLPALTYH-----VTOPK-----DGFCEVC 406
DB 342 TYSPSLVQ-LVAKITPEKVKCFRLC-GNRRARAVHDAVAIVPSPEDMAENQSFQNGC 399
QY 407 KKLVGYLDRNLEKNSTKOEILALKEKGSFLPDYOKOCDOFVAEYEPVILIVEVMP 466
DB 400 KRLITVSSHNLSEKSTKRDILVAFKGGCSILPLPYMIQCKHFTVQYEPVILISLKMMDP 459
QY 467 SFVCLKIGACPSAHKRLGTETKICWPSYWCNTETAAQCNVHECKRHVW 517
DB 460 VAVCKKVGACHGPRTPILGTQCALGPSPWCRSGQAALKCAVHCGQHVM 510

```

## RESULT 11

```

US-09-833-245-903
; Sequence 903, Application US/09833245
; Publication No. US20040010134A1
; GENERAL INFORMATION:
; APPLICANT: Human Genome Sciences, Inc.
; TITLE OF INVENTION: Albumin Fusion Proteins
; FILE REFERENCE: PFS46PCT
; CURRENT APPLICATION NUMBER: US/09/833,245
; CURRENT FILING DATE: 2001-04-12
; PRIOR APPLICATION NUMBER: 60/229,358
; PRIOR FILING DATE: 2000-04-12
; PRIOR APPLICATION NUMBER: 60/256,931
; PRIOR FILING DATE: 2000-12-21
; PRIOR APPLICATION NUMBER: 60/199,384
; NUMBER OF SEQ ID NOS: 2267
; SOFTWARE: Patent In Ver. 2.1
; SEQ ID NO 903
; LENGTH: 531
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-833-245-903

```

```

Query Match 40.9%; Score 1127.5; DB 11; Length 531;
Best Local Similarity 42.9%; Pred. No. 1.5e-91;
Matches 228; Conservative 99; Mismatches 161; Indels 43; Gaps 11;
QY 1 LASLIGALAGPVLGLKECTRGSAVMCONVKTASDCGAKHCLQTYWKNKPTVSLACDTC 60
DB 8 LPSLIGATRASPSGPECAKSTVMQODIQOTARCAVGTCCGAANNKPTASLPCDVC 67
QY 61 KDVTAAAGDMLKDNATEEELIVYEKTCMDLPRKNNASCKEIVDSYLPVILDIIGKEM- 119
DB 68 QDIAAAGNGINPATESDITLALVMKTCFWLPQESSAGCKMMVDAHSSAILSLMARGAD 127

```

```

QY 120 SRPGEVSALNLCSLQKHLAEINHOQKLESNKIPELDMTEVAVAPPMANIPILLYPODGP 179
DB 128 SAPAQVCTALSLCEPLQRHLATL-----RPLSKEDTEBAVAPPMANIPILFFHROAP 179
QY 180 RSKPQPKNDGVQDCDIQMTDITQAVRTNSTFVQALVEHVEKBCDRLPGMADICKNYI 239
DB 180 -----EGALCDCCVRQVSRLOEAVRSNLTADL---NIQRCESLPGIHLAVLCNLY 228
QY 240 SQYSEIALQMMHMQPKKEICALVGFCDVEKEMQTLVPAK---VASKNVIPLBELVEPI 296
DB 229 FQFVFPADQALRLPPELCEKRGKGFCEELG-----APARLTQVAMDGVPSELGLPR 281
QY 297 KKHVEPAKSDVYCEVCEFLVKEVTKLIDNNKTEKEILDAFDMKCSKLPKSLSEECQEVND 356
DB 282 KQSEMOMKAGVTCVCECMNVVQKLDHWMNSSELMITHALERCSVMPASITKECIIIVD 341
QY 357 TYGSSILSLLEESPELVCSMLHLCSGTRLPALTYH-----VTOPK-----DGFCEVC 406
DB 342 TYSPSLVQ-LVAKITPEKVKCFRLC-GNRRARAVHDAVAIVPSPEDMAENQSFQNGC 399
QY 407 KKLVGYLDRNLEKNSTKOEILALKEKGSFLPDYOKOCDOFVAEYEPVILIVEVMP 466
DB 400 KRLITVSSHNLSEKSTKRDILVAFKGGCSILPLPYMIQCKHFTVQYEPVILISLKMMDP 459
QY 467 SFVCLKIGACPSAHKRLGTETKICWPSYWCNTETAAQCNVHECKRHVW 517
DB 460 VAVCKKVGACHGPRTPILGTQCALGPSPWCRSGQAALKCAVHCGQHVM 509

```

## RESULT 12

```

US-10-043-487-340
; Sequence 340, Application US/10043487
; Publication No. US20030055220A1
; GENERAL INFORMATION:
; APPLICANT: HYBRIGENICS
; APPLICANT: Pierre, LEGRAIN
; TITLE OF INVENTION: Protein-protein interactions between Shigella flexneri polypeptide
; FILE REFERENCE: B4778A
; CURRENT APPLICATION NUMBER: US/10/043,487
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/261,130
; PRIOR FILING DATE: 2001-01-12
; NUMBER OF SEQ ID NOS: 561
; SOFTWARE: Patent In version 3.1
; SEQ ID NO 340
; LENGTH: 209
; TYPE: PRT
; ORGANISM: Shigella flexneri
US-10-043-487-340

```

```

Query Match 39.6%; Score 1093; DB 14; Length 209;
Best Local Similarity 100.0%; Pred. No. 4.ee-89;
Matches 209; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 145 OKLESNKIPELDMTEVAVAPPMANIPILLYPODGPSPKQPKNDGVQDCDIQMTDITQ 204
DB 1 OKLESNKIPELDMTEVAVAPPMANIPILLYPODGPSPKQPKNDGVQDCDIQMTDITQ 60
QY 205 AVRTNSTFVQALVEHVEKBCDRLPGMADICKNYISQYSEIALQMMHMQPKKEICALVGF 264
DB 61 AVRTNSTFVQALVEHVEKBCDRLPGMADICKNYISQYSEIALQMMHMQPKKEICALVGF 120
QY 265 CDEVEKEMQTLVPAKASKNVILPALBELVEPIKHEVPAKSDVYCEVCEFLVKEVTKLID 324
DB 121 CDEVEKEMQTLVPAKASKNVILPALBELVEPIKHEVPAKSDVYCEVCEFLVKEVTKLID 180
QY 325 NNKTEKEILDAFDMKCSKLPKSLSEECQOE 353
DB 181 NNKTEKEILDAFDMKCSKLPKSLSEECQOE 209

```

## RESULT 13

US-10-332-426-8  
; Sequence 8, Application US/10332426  
; Publication No. US20040029136A1  
; GENERAL INFORMATION:  
; APPLICANT: INCYTE GENOMICS, INC.; TANG, Y. Tom;  
; APPLICANT: AZIMZAI, Valida; DAS, Debopriya;  
; APPLICANT: THORNTON, Michael; LU, Dying; Aina M.;  
; APPLICANT: TRIBOUTLEY, Catherine M.; YUE, Henry;  
; APPLICANT: GANDHI, Ameena R.; CHAMLA, Narinder K.;  
; APPLICANT: KHAN, Farrah A.; LU, Yan;  
; APPLICANT: YAO, Monique G.; HAFALIA, April J. A.;  
; APPLICANT: ELIOT, Vicki S.; ARIZU, Chandra S.;  
; APPLICANT: LAL, Preeti; RAMKUMAR, Jayalaxmi;  
; APPLICANT: NGUYEN, Damiel B.; BAUGHN, Mariah R.  
; TITLE OF INVENTION: LIPID METABOLISM MOLECULES  
; FILE REFERENCE: PI-0152 USN  
; CURRENT APPLICATION NUMBER: US/10/332,426  
; CURRENT FILING DATE: 2003-01-06  
; PRIOR APPLICATION NUMBER: US 60/216,803  
; PRIOR FILING DATE: 2000-07-07  
; PRIOR APPLICATION NUMBER: US 60/216,801  
; PRIOR FILING DATE: 2000-07-07  
; PRIOR APPLICATION NUMBER: US 60/218,233  
; PRIOR FILING DATE: 2000-07-14  
; PRIOR APPLICATION NUMBER: US 60/220,046  
; PRIOR FILING DATE: 2000-07-21  
; PRIOR APPLICATION NUMBER: US 60/220,739  
; PRIOR FILING DATE: 2000-07-26  
; PRIOR APPLICATION NUMBER: US 60/222,824  
; PRIOR FILING DATE: 2000-08-04  
; NUMBER OF SEQ. ID NOS: 16  
; SOFTWARE: PERL Program  
; SEQ. ID NO 8  
; LENGTH: 362  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: misc feature  
; OTHER INFORMATION: Incyte ID No. US20040029136A1 7473574CD1  
US-10-332-426-8  
Query Match 29.8%; Score 822.5; DB 12; Length 362;  
Best Local Similarity 33.0%; Pred. No. 1.4e-64;  
Matches 174; Conservative 55; Mismatches 104; Indels 195; Gaps 7;  
QY 1 LASLIGALAGVILGKECTRGSAVWQNVKTASDCAVHCLQTVWKNKPTVYSLPCDIC 60  
DB 8 LPSLIGATRASPTSGPQCAKSGSTWCQDLQTAARCAVGYCGAVWKNKPTAKSLPCDVC 67  
QY 61 KQVVTAGGMLKDNATFEELIYLEKTCWMLPKPMNSASCKEIVDSYLFVILDIKGEK- 119  
DB 68 QDIAAAGNGINPDATESDILALVMKTCWMLPSQESSACKWVDAHSSAILISMARGAD 127  
QY 120 SRPGEVCSALNLCESLQKHLAEINHOXOLESNKIPELDTEVAVAPPMANIPILLVPODGP 179  
DB 128 SARPAVCTRLSTCEPQRLHATL- RPLSEDTFEAVAPPMANGPLTFHRQAP 179  
QY 180 RSKPQKNGVCDQCIQWVTDIQTAVRTNSTFVQALVEHVKESCDRLGPMADICKNYI 239  
DB 180 -----EGALCQDCVRO----- 190  
QY 240 SGYSEIATQMMHMQPKKEICALVGFCEVEMPMQTLVPAKVASKNVILPALSLVEPIKHA 299  
DB 191 ----- 190  
QY 300 EYPAKSDVYCEVFLVKEVTKLIDNNKTEKEILDADFQWCKSLPKSLSEBQEVVDYVG 359  
DB 191 ----- 190  
QY 360 SSILSILEEVSPELVCSMLHLCGSTRLPALTVAH-----VTQPK-----DGFCEVCKKL 409  
DB 191 ----- 190  
QY 191 -----LVAKITPEKVKCFIRLC-GNRRBARAVHDAVIAIVSPBEMDAENQSGFCNGCKRL 243

QY 410 VGYLDRLNEKSTKOEIILAELKSGSFLPDPYQKQDQFVAEYEPVLIETLVEMDPSFV 469  
DB 244 LTVSSHNIESTKSTKDIILVAKRGCSILPLPYMIOCKAFVYQYEVVLIESTKOMMDPAV 303  
QY 470 CLKIGACPSAHKPLIGTEKCIWGPSPYWCQNTETAQCAVNAHEKRRHW 517  
DB 304 CKKVACGHPRPPLIGTDQCALGSPFCWCRSGEAAATLCAVAVHQCQGHV 351  
RESULT 14  
US-10-267-502-385  
; Sequence 385, Application US/10267502  
; Publication No. US20040071700A1  
; GENERAL INFORMATION:  
; APPLICANT: Kim, Jaeseob  
; APPLICANT: Galant, Ron  
; TITLE OF INVENTION: Obesity Linked Genes  
; FILE REFERENCE: LSD-07416  
; CURRENT APPLICATION NUMBER: US/10/267,502  
; CURRENT FILING DATE: 2003-01-27  
; NUMBER OF SEQ. ID NOS: 439  
; SOFTWARE: PatentIn version 3.2  
; SEQ. ID NO 385  
; LENGTH: 953  
; TYPE: PRT  
; ORGANISM: Drosophila melanogaster  
US-10-267-502-385  
Query Match 19.9%; Score 549.5; DB 12; Length 953;  
Best Local Similarity 23.2%; Pred. No. 1.2e-39;  
Matches 144; Conservative 105; Mismatches 212; Indels 161; Gaps 18;  
QY 8 ALAGPVILGKECTRGSAVWQNVKTASDCAVHCLQTVWKNK---PTVYSLPCDICQDV 64  
DB 20 AAATPILSSKCTWGPVYCGNFSNCKECPATRHCIQTIVWETQKVPVDTDSICTICDMV 79  
QY 65 TAAQDMTKDNATFEELIYLEKTCWMLPKPMNSASCKEIVDSYLFVILDIKGEKSRGGE 124  
DB 80 TQARQQLNSNOTEEELKEVFESGCKLIPKPIQKECIVADLPBELVLAALQNN-PDQ 138  
QY 125 VCSALNLCES-----LQKH-----LAEIN 143  
DB 139 VCSVAGLGNARSARIDELVKNGIQAGLDVTQNBDSSEFTELAMQNPQSLSCGNLNLSTLM 198  
QY 144 HOKOLESNKIPELDMTEVAV-----APPMANIPILL----- 173  
DB 199 HSKFPAATRD-----DWEETMLHMGSLSSFSDACANIVLTVFNDIYDHSKILTTDAVCHV 255  
QY 174 -----YPOGPRKQPK-----DNGD-----VQDQCIQWVTDIQTAVRTNSTFVQALV 217  
DB 256 SGVCASRYHQHEEKQPEBALVALDAGDDIPCELCQQLVKHLRDVLVANNTETEKQVME 315  
QY 218 EHVKEECRLRPGMADICKNYISQYSEIAIQ-MMMHMQPKKEICALVGFCEDE-----VKEM 271  
DB 316 GECKQ-----SKGPFDELSLIVDQYHYIYETIVSKLDANAGCCMIGICQGRNASMSMDV 370  
QY 272 PMQTLVPAKVASKNVILAEU-----VEPIKKE-----VP----- 302  
DB 371 PIMPLLP-----VIEPAQVKITTEKLEKKEKQGLASBPFSQOETLMDQLPIDHLMG 423  
QY 303 -----AKSDVYCEVCEFLVKEVTKLIDNNKTEKEIILDAFDKMSCKLPKSLSEBQEVV 355  
DB 424 AANPGLAVEGELCTLCYMHAFIQTETATSTDEIHGTVENICAKLPBGVAGQCRNFV 483  
QY 356 DTVGSSISILLEEVSPELVCSMLHLCGSTRLPALTVAHVP-----KDGFCFVCKKL 409  
DB 484 EYMGDAVIALVQGINPDPVCPIMQMCNLPKKEDEVAVRPNQPSDDQDPPTGCLCLFA 543  
QY 410 VGYLDRLNEKSTKOEIILAELKSGSFLPDPYQKQDQFVAEYEPVLIETLVEMDPSFV 469  
DB 544 VEOAQMKTRDNKSKONIKVNLGCLSHLPNEIKECCVDFVVTYNSNELLMDITDFKPOEI 603  
QY 470 CLKIGACPSAHKPLIGTEKCIW 491



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

## OM protein - protein search, using sw model

Run on: May 5, 2004, 13:26:23 ; Search time 20.661 Seconds

(without alignments)  
2411.659 Million cell updates/secTitle: US-09-743-684a-1\_COPY\_7\_524  
score greater than or equal to the score of the result being printed,  
Sequence: 1 LASLIGALAGPVGLKECT.....NTETAAQCNAYEHCRRHWN 518Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 283366 seqs, 96191526 residues

Total number of hits satisfying chosen parameters: 283366

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : PIR 78:\*

1: pir1:.\*  
2: pir2:.\*  
3: pir3:.\*  
4: pir4:.\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2747.5	99.6	527	1	SAHUP
2	1974	71.5	554	1	A28716
3	1915.5	69.4	557	2	JH0604
4	511	18.5	965	2	T00207
5	389	14.1	79	2	A49475
6	372	13.5	376	2	S02766
7	343	12.4	80	2	S21770
8	339.5	12.3	381	1	LNHUB
9	305	11.1	363	2	A29072
10	293	10.6	369	2	I46531
11	261	9.5	81	2	A32026
12	250	9.1	370	1	LNRBB
13	213	7.7	213	2	T46069
14	203.5	7.4	402	2	T148201
15	176.5	6.4	412	2	T15677
16	173.5	6.3	513	2	T09739
17	171.5	6.2	506	2	F86253
18	165.5	6.0	506	2	T07915
19	161.5	5.9	513	2	T11686
20	161	5.8	428	2	S47096
21	161	5.8	474	2	T12049
22	158.5	5.7	508	2	S19697
23	156	5.7	314	2	T15674
24	155.5	5.6	292	2	T14446
25	155.5	5.6	322	2	S41400
26	154	5.6	496	2	JU0732
27	153.5	5.6	409	2	JC7272
28	145.5	5.3	205	2	B89567
29	143.5	5.2	1175	2	S52417

30 142.5 5.2 509 2 S66516  
31 142.5 5.2 509 2 S49349  
32 142.5 5.2 1948 2 S00485  
33 141 5.1 508 2 DB5056  
34 134 4.9 1927 2 G64585  
35 132.5 4.8 433 2 E96649  
36 127.5 4.6 280 2 PC4080  
37 127.5 4.6 1142 2 A45031  
38 126.5 4.6 195 2 T15676  
39 126.5 4.6 1819 2 A71928  
40 126 4.6 506 2 S71591  
41 123.5 4.5 370 2 E96502  
42 116.5 4.2 652 2 F85017  
43 116.5 4.2 707 2 T26218  
44 116 4.2 975 2 T59422  
45 116 4.2 1357 2 B96696

oryzasin (EC 3.4.2  
cyprosin (EC 3.4.2  
gene 11-1 protein  
probable aspartic  
cag pathogenicity  
hypothetical prote  
aspartic proteinas  
cysteine-rich fibr  
hypothetical prote  
cag island protein  
aspartic proteinas  
hypothetical prote  
probable CHP-rich  
hypothetical prote  
resc8 - rat (fragm  
protein FIN21.4 [1

## ALIGNMENTS

## RESULT 1

## SAHUP

saposin precursor [validated] - human  
N:Alternate names: cerebroside sulfate activator protein; co-beta-glucosidase; component  
ein (SAP); sphingolipid activator protein A2; sulfatide sulfatase activator protein  
N:Contains: prosaposin; saposin A; saposin B; saposin C; saposin D  
C:Species: Homo sapiens (man)  
C:Date: 30-Jun-1992 #sequence revision 17-Nov-1995 #text change 08-Dec-2000  
C:Accession: JX0061; A57368; A42003; B42003; C42003; A30367; S34740; S36140; S36  
0226; I37265; I37264

R:Nakano, T.; Sandhoff, K.; Stuenkel, J.; Christomanou, H.; Suzuki, K.  
J. Biochem. 105, 152-154, 1989  
A:Title: Structure of full-length cDNA coding for sulfatide activator, a Co-beta-glucosi-  
A:Reference number: JX0061; MUID:89255151; PMID:2498298  
A:Accession: JX0061

A:Molecule type: mRNA  
A:Residues: 1-527 <NMK>  
A:Cross-references: GB:D00422; NID:g220063; PIDN:BAA00321.1; PID:g220064  
A>Note: alternative splice form 1  
A:Accession: A57368

A:Molecule type: mRNA  
A:Residues: 1-259, 263-527 <NA2>  
A:Cross-references: GB:J03015; GB:J03086; NID:g337755; PIDN:AAB59494.1; PID:g337756  
A>Note: alternative splice form 2  
R:Rotman, E.G.; Scheinker, V.; Grabowski, G.A.  
Genomics 13, 312-318, 1992

A:Title: Structure and evolution of the human prosaposin chromosomal gene.  
A:Reference number: A42003; MUID:92307663; PMID:1612590  
A:Accession: A42003

A:Molecule type: DNA  
A:Residues: 50-140 <ROR>  
A:Cross-references: GB:M6181  
A>Note: sequence extracted from NCBI backbone (NCBIN:107235, NCBIIP:107236)  
A:Accession: B42003

A:Molecule type: DNA  
A:Residues: 185-259, 263-276 <RO2>  
A>Note: sequence extracted from NCBI backbone (NCBIN:107235, NCBIIP:107237)  
A:Accession: C42003

A:Molecule type: DNA  
A:Residues: 305-393 <RO3>  
A>Note: sequence extracted from NCBI backbone (NCBIN:107235, NCBIIP:107238); sequence inc  
A:Accession: D42003

A:Molecule type: DNA  
A:Residues: 399-487 <RO4>  
A>Note: sequence extracted from NCBI backbone (NCBIN:107235, NCBIIP:107239); sequence inc  
R:Rotman, E.G.; Grabowski, G.A.  
Genomics 5, 486-492, 1989

A:Title: Molecular cloning of a human co-beta-glucosidase cDNA: evidence that four sphin-  
A:Reference number: A30367; MUID:90129043; PMID:2515150  
A:Accession: A30367  
A:Molecule type: mRNA  
A:Residues: 1-259, 263-527 <RO5>

A/Cross-references: GB:J03077; NID:9183230; PIDN:AAA52560.1; PID:9183231  
A/Note: alternative splice form 2  
R/Hirata, M.; O'Brien, J.S.; Kishimoto, Y.; Galdzicka, M.; Fluharty, A.L.; Gims, E.I.; Arch. Biochem. Biophys. 304, 110-116, 1993  
A/Title: Isolation, characterization, and proteolysis of human prosaposin, the precursor of  
A/Accession: S34740  
A/Molecule type: protein  
A/Residues: 17-24;165-172;180-189;301-305 <HR>  
R/Tymela, J.; Palmer, D.N.; Baumann, M.; Halcia, M. FEBS Lett. 330, 8-12, 1993  
A/Title: Storage of saposin A and D in infantile neuronal ceroid-lipofuscinosis.  
A/Reference number: S36140; MUID:93380576; PMID:8370464  
A/Accession: S36140  
A/Molecule type: protein  
A/Residues: 'XX', 62, 'X', 64-65, 'X', 67-79, 'X', 81-84 <TV>  
A/Note: saposin A  
A/Accession: S36141  
A/Molecule type: protein  
A/Residues: 'XXX', 413-414, 'X', 416-428, 'X', 430-434 <TV>  
A/Note: saposin D  
R/Holtschmidt, H.; Sandhoff, K.; Kwon, H.Y.; Harzer, K.; Nakano, T.; Suzuki, K. J. Biol. Chem. 266, 7556-7560, 1991  
A/Title: Sulfatide activator protein. Alternative splicing that generates three mRNAs at  
A/Reference number: S36988; MUID:91210267; PMID:2019586  
A/Accession: S36988  
A/Status: nucleic acid sequence not shown; translation not shown  
A/Molecule type: mRNA  
A/Residues: 1-240, 'S', 242-252 <HOL>  
A/Cross-references: EMBL:M60255; NID:9337759; PIDN:AAA6594.1; PID:9337760  
A/Note: the nucleotide sequence was submitted to the EMBL Data Library, June 1991  
A/Title: cerebroside sulfate activator protein mutant MU-9; corresponds to alternative sp  
A/Accession: S36989  
A/Status: nucleic acid sequence not shown; translation not shown  
A/Molecule type: mRNA  
A/Residues: 1-240, 'S', 242-252, 263-527 <HO2>  
A/Cross-references: EMBL:M60257; NID:9337764; PIDN:AAA6595.1; PID:9337765  
A/Note: the nucleotide sequence was submitted to the EMBL Data Library, June 1991  
A/Title: cerebroside sulfate activator protein mutant MU-0; corresponds to alternative sp  
A/Accession: S36990  
A/Status: nucleic acid sequence not shown; translation not shown  
A/Molecule type: mRNA  
A/Residues: 1-240, 'S', 242-252, 261-527 <HO3>  
A/Cross-references: EMBL:M60258; NID:9337766; PIDN:AAA6596.1; PID:9337767  
A/Note: the nucleotide sequence was submitted to the EMBL Data Library, June 1991  
A/Title: cerebroside sulfate activator protein mutant MU-6; corresponds to alternative sp  
R/Koch, K.; Hienho, T.; Sano, A.; Kakimoto, Y. Biochem. Biophys. Res. Commun. 181, 286-292, 1991  
A/Title: Isolation and characterization of prosaposin from human milk.  
A/Reference number: P80330; MUID:92068206; PMID:1958198  
A/Accession: P80330  
A/Molecule type: protein  
A/Residues: 17-24, 'X', 26 <KON>  
A/Experimental source: milk  
R/Kretz, K.A.; Carson, G.S.; Morimoto, S.; Kishimoto, Y.; Fluharty, A.L.; O'Brien, J.S. Proc. Natl. Acad. Sci. U.S.A. 87, 2541-2544, 1990  
A/Title: Characterization of a mutation in a family with saposin B deficiency: a glycosy  
A/Reference number: A35985; MUID:90207231; PMID:2320574  
A/Accession: A35985  
A/Molecule type: mRNA  
A/Residues: 213-221 <KRE>  
A/Cross-references: GB:M32221  
A/Accession: B35985  
A/Status: nucleic acid sequence not shown; translation not shown  
A/Molecule type: mRNA  
A/Residues: 1-259, 263-527 <KR2>  
A/Cross-references: GB:M32221; NID:9337761; PIDN:AAA60303.1; PID:9337762  
A/Experimental source: lymphoblast  
A/Accession: C35985  
A/Molecule type: mRNA  
A/Residues: 213-216, 'I', 218-221 <KR3>  
A/Note: sequence from patients with activator-deficient metachromatic leukodystrophy; th  
R/Pieret, W.; Schubert, J.; Machleidt, W.; Meyer, H.E.; Sandhoff, K.

Eur. J. Biochem. 192, 709-714, 1990  
A/Title: The complete amino-acid sequences of human ganglioside GM2 activator protein an  
A/Reference number: S13195; MUID:9106165; PMID:2209618  
A/Accession: S13196  
A/Molecule type: protein  
A/Residues: 195-259, 263-277 <FUE>  
R/Morimoto, S.; Martin, B.M.; Yamamoto, Y.; Kretz, K.A.; O'Brien, J.S.; Kishimoto, Y. Proc. Natl. Acad. Sci. U.S.A. 86, 3389-3393, 1989  
A/Title: Saposin A: second cerebrosidase activator protein.  
A/Reference number: A32784; MUID:89240739; PMID:2717620  
A/Accession: A32784  
A/Molecule type: protein  
A/Residues: 60-84;86-107;109-119;125-134 <MOR>  
R/O'Brien, J.S.; Kretz, K.A.; Dewji, N.; Wenger, D.A.; Each, F.; Fluharty, A.L. Science 241, 1098-1101, 1988  
A/Title: Coding of two sphingolipid activator proteins (SAP-1 and SAP-2) by same genetic  
A/Reference number: A41240; MUID:88321660; PMID:2842863  
A/Accession: A41240  
A/Molecule type: mRNA  
A/Residues: 'GSSR', 18-259, 263-299, 'D', 301-302, 'D', 304-527 <OAB>  
A/Cross-references: GB:J03086  
R/Dewji, N.N.; Wenger, D.A.; O'Brien, J.S. Proc. Natl. Acad. Sci. U.S.A. 84, 8552-8556, 1987  
A/Title: Nucleotide sequence of cloned cDNA for human sphingolipid activator protein 1 p  
A/Reference number: S02289; MUID:88068647; PMID:2825202  
A/Accession: S02289  
A/Status: significant sequence differences  
A/Molecule type: mRNA  
A/Cross-references: EMBL:J03015  
A/Note: this sequence corrected by A41240  
R/Kleinhardt, T.; Christomanou, H.; Braunitzer, G. Biol. Chem. Hoppe-Seyler 369, 1361-1365, 1988  
A/Title: Complete amino-acid sequence of the naturally occurring A(2) activator protein  
A/Reference number: S02028; MUID:89207118; PMID:3242555  
A/Accession: S02028  
A/Molecule type: protein  
A/Residues: 195-259, 263-276 <KLE>  
R/Pieret, W.; Machleidt, W.; Sandhoff, K. Biol. Chem. Hoppe-Seyler 369, 317-328, 1988  
A/Title: The precursor of sulfatide activator protein is processed to three different pr  
A/Reference number: S00813; MUID:89000190; PMID:3048308  
A/Accession: S00813  
A/Molecule type: protein  
A/Residues: 410-487 <PU2>  
R/Kleinhardt, T.; Christomanou, H.; Braunitzer, G. Biol. Chem. Hoppe-Seyler 368, 1571-1578, 1987  
A/Title: Complete amino-acid sequence and carbohydrate content of the naturally occurring  
A/Reference number: S00226; MUID:88163077; PMID:3442600  
A/Accession: S00226  
A/Molecule type: protein  
A/Residues: 314-393 <KL2>  
R/Vaccaro, A.M.; Salvio, R.; Barca, A.; Tatti, M.; Ciaffoni, F.; Maras, B.; Siciliano, U. Biol. Chem. 270, 9953-9960, 1995  
A/Title: Structural analysis of saposin C and B. Complete localization of disulfide bridg  
A/Reference number: A57297; MUID:95247790; PMID:7730378  
A/Contents: annotation; disulfide bonds; glycosylation  
R/Holtschmidt, H.; Sandhoff, K.; Pieret, W.; Kwon, H.Y.; Schnabel, D.; Suzuki, K. FEBS Lett. 280, 267-270, 1991  
A/Title: The organization of the gene for the human cerebroside sulfate activator protei  
A/Reference number: I37264; MUID:91192146; PMID:2013321  
A/Accession: I37265  
A/Status: translated from GB/EMBL/DBJ  
A/Molecule type: DNA  
A/Residues: 59-125 <RES>  
A/Cross-references: EMBL:X57107; NID:930234; PIDN:CAA40391.1; PID:930235  
A/Accession: I37264  
A/Status: translated from GB/EMBL/DBJ  
A/Molecule type: DNA  
A/Residues: 307-516 <RE2>  
A/Cross-references: EMBL:X57108; NID:930232; PIDN:CAA40392.1; PID:91565257  
A/Note: sequence revised relative to PID:930233 (corrected coding region)  
C/Genetics:



A:Gene: GDB:PSAP; GLBA  
A:Cross-references: GDB:120366; OMIM:176801  
A:Map position: 10q22.1-10q22.1  
A:Introns: 83/3; 338/3; 401/1; 453/3; 480/3  
A>Note: defects in this gene may cause variant Gaucher disease, variant Tay-Sachs disease  
A>Note: list of introns is incomplete

Query Match 99.6%; Score 2747.5; DB 1; Length 527;  
Best Local Similarity 99.4%; Pred. No. 176-173; Indels 3; Gaps 1;  
Matches 518; Conservative 0; Mismatches 0; Indels 3; Gaps 1;

```
QY 1 LASLGAALAGPVGLKCTRGSAVWCONVKTASDCAVKHCLQTVMNKPVTYSLPCIC 60
DB 7 LASLGAALAGPVGLKCTRGSAVWCONVKTASDCAVKHCLQTVMNKPVTYSLPCIC 66
QY 61 KDVTAAAGDMLKDNATBEELIVLEKTCMDLPKPMNSASCKEIVDSYLPVLIIDIKGMS 120
DB 67 KDVTAAAGDMLKDNATBEELIVLEKTCMDLPKPMNSASCKEIVDSYLPVLIIDIKGMS 126
QY 121 RRGVCSALNLCESLOKHLAELNHOKLESNKIPRLDTEVVAPEPMANIPLLIYPODGR 180
DB 127 RRGVCSALNLCESLOKHLAELNHOKLESNKIPRLDTEVVAPEPMANIPLLIYPODGR 186
QY 181 SKPOKNDGVCOQDICIQWVTDIQTAVRTNSTFVQALVHVKEECRLGPGMADICKNYIS 240
DB 187 SKPOKNDGVCOQDICIQWVTDIQTAVRTNSTFVQALVHVKEECRLGPGMADICKNYIS 246
QY 241 QYSEIATOMMMHMOOQPKKEICLVGFCDVEKEMPMQTLVPAKVASKNVIPLALVEPEIK 297
DB 247 QYSEIATOMMMHMOOQPKKEICLVGFCDVEKEMPMQTLVPAKVASKNVIPLALVEPEIK 306
QY 298 KHEVPAKSDVYCEVCEFLVKEVTKLIDNNKTEKEILDADFMCCKLPKSLSECEQVDT 357
DB 307 KHEVPAKSDVYCEVCEFLVKEVTKLIDNNKTEKEILDADFMCCKLPKSLSECEQVDT 366
QY 358 YGSSILSTLLEBVSPELVCSMLHLCSGTRLPALTVHTVQPKDGFCEVCKLVGLDRL 417
DB 367 YGSSILSTLLEBVSPELVCSMLHLCSGTRLPALTVHTVQPKDGFCEVCKLVGLDRL 426
QY 418 EKNSTKOEILALAEKGSFLLPDPYKQCDQFVAEYEPVILIEIVWMDPSFVCLKIGACP 477
DB 427 EKNSTKOEILALAEKGSFLLPDPYKQCDQFVAEYEPVILIEIVWMDPSFVCLKIGACP 486
QY 478 SAHKPLIGTEKCIWGPSTWCONTEPAACNAVEHCKRHVN 518
DB 487 SAHKPLIGTEKCIWGPSTWCONTEPAACNAVEHCKRHVN 527
```

## RESULT 2

A28716

saposin precursor - rat

N:Alternate names: cerebroside sulfate activator protein; co-beta-glucosidase; component  
ein (SAP); sphingolipid activator protein A2; sulfated glycoprotein 1; sulfatide sulfate  
N:Contains: prosaposin; saposin A; saposin B; saposin C; saposin D  
C:Species: Rattus norvegicus (Norway rat)  
C:Date: 10-Sep-1999 #sequence\_revision 10-Sep-1999 #text\_change 10-Sep-1999  
C:Accession: A28716

R:Collard, M.W.; Sylvestre, S.R.; Tsunrta, J.K.; Griswold, M.D.  
Biochemistry 27, 4557-4564, 1988

A:Title: Biosynthesis and molecular cloning of sulfated glycoprotein 1 secreted by rat B  
A:Reference number: A28716; MUID:89006447; PMID:3048385

A:Accession: A28716

A:Molecule type: mRNA

A:Residues: 1-554 &lt;CDS&gt;

A:Cross-references: GB:M9936; NID:9206904; PID:AAA42136.1; PID:9206905  
A>Note: parts of this sequence, including the amino end of the mature protein, were de

C:Function:  
A:Description: saposin bind sphingolipids, form hydrophilic complexes and make them acc

A:Pathway: sphingolipid catabolism

A>Note: saposin A and C (SAP-2) activate hydrolysis of glucocerebroside by beta-glucosy

A>Note: saposin B (SAP-1) activates hydrolysis of galactocerebroside sulfate by arylsul

A>Note: saposin D activates hydrolysis of sphingomyelin by sphingomyelin phosphodiester

C:Superfamily: saposin; saposin repeat homology

C:Keywords: alternative splicing; glycoprotein; lysosomal storage disease; lysosome; sph

F:1-16/Domain: signal sequence #status predicted <SIG>

F:17-554/Product: prosaposin #status predicted <PRO>

F:55-148/Domain: saposin repeat homology <SAP1>

F:160-143/Product: saposin A #status predicted <SAPA>

F:189-280/Domain: saposin repeat homology <SAP2>

F:194-273/Product: saposin B #status predicted <SAB1>

F:306-397/Domain: saposin repeat homology <SAP3>

F:310-389/Product: saposin C #status predicted <SAPC>

F:431-522/Domain: saposin repeat homology <SAP4>

F:437-514/Product: saposin D #status predicted <SAPD>

F:63-138,66-132,94-106,439-512,442-506,470-481/disulfide bonds: #status predicted

F:80,214,331,456/Binding site: carbohydrate (Asn) (covalent) #status predicted

F:197-270,200-264,229-240,314-387,317-381,345-356/disulfide bonds: #status predicted

Query Match 71.5%; Score 1974; DB 1; Length 554;  
Best Local Similarity 66.1%; Pred. No. 166-122;  
Matches 363; Conservative 77; Mismatches 77; Indels 32; Gaps 3;

```
QY 1 LASLGAALAGPVGLKCTRGSAVWCONVKTASDCAVKHCLQTVMNKPVTYSLPCIC 60
DB 7 LASLGAALAGPVGLKCTRGSAVWCONVKTASDCAVKHCLQTVMNKPVTYSLPCIC 66
QY 61 KDVTAAAGDMLKDNATBEELIVLEKTCMDLPKPMNSASCKEIVDSYLPVLIIDIKGMS 120
DB 67 KDVTAAAGDMLKDNATBEELIVLEKTCMDLPKPMNSASCKEIVDSYLPVLIIDIKGMS 126
QY 121 RRGVCSALNLCESLOKHLAELNHOKLESNKIPRLDTEVVAPEPMANIPLLIYPODGR 180
DB 127 RRGVCSALNLCESLOKHLAELNHOKLESNKIPRLDTEVVAPEPMANIPLLIYPODGR 185
QY 181 SKPOKNDGVCOQDICIQWVTDIQTAVRTNSTFVQALVHVKEECRLGPGMADICKNYIS 240
DB 186 SKPOKNDGVCOQDICIQWVTDIQTAVRTNSTFVQALVHVKEECRLGPGMADICKNYIS 245
QY 241 QYSEIATOMMMHMOOQPKKEICLVGFCDVEKEMPMQTLVPAKVASKNVIPLALVEPEIK 300
DB 246 QYSEIATOMMMHMOOQPKKEICLVGFCDVEKEMPMQTLVPAKVASKNVIPLALVEPEIK 305
QY 301 VPAKSDVYCEVCEFLVKEVTKLIDNNKTEKEILDADFMCCKLPKSLSECEQVDTYGS 360
DB 306 VPAKSDVYCEVCEFLVKEVTKLIDNNKTEKEILDADFMCCKLPKSLSECEQVDTYGS 365
QY 361 SILSTLLEBVSPELVCSMLHLCSG-----TRLP-----A 389
DB 366 SILSTLLEBVSPELVCSMLHLCSG-----TRLP-----A 389
QY 390 LTVHTQPKDGFCEVCKLVGLDRLNLEKNSKOEILALAEKGSFLLPDPYKQCDQFV 449
DB 426 LRAHVPKNGGFCFCEVCKLVGLDRLNLEKNSKOEILALAEKGSFLLPDPYKQCDQFV 485
QY 450 AEYEPVILIEIVWMDPSFVCLKIGACPSAHKPLIGTEKCIWGPSTWCONTEPAACNAV 509
DB 486 AEYEPVILIEIVWMDPSFVCLKIGACPSAHKPLIGTEKCIWGPSTWCONTEPAACNAV 545
QY 510 EHCKRHVN 518
DB 546 EHCKRHVN 554
```

## RESULT 3

JH0604

saposin precursor - mouse

N:Alternate names: cerebroside sulfate activator protein; co-beta-glucosidase; component  
ein (SAP); sphingolipid activator protein A2; sulfated glycoprotein 1; sulfatide sulfate  
N:Contains: prosaposin; saposin A; saposin B; saposin C; saposin D  
C:Species: Mus musculus (house mouse)

C:Date: 10-Sep-1999 #sequence\_revision 10-Sep-1999 #text\_change 10-Sep-1999

C:Accession: JH0604

R:Tajima, M.; Sakiyama, T.; Endo, H.; Kitagawa, T.

Biochem. Biophys. Res. Commun. 184, 1266-1272, 1992

A:Title: The primary structure of mouse saposin.

A:Reference number: JH0604; MUID:92272718; PMID:1590788

A:Accession: JH0604  
 A:Molecule type: mRNA  
 A:Residues: 1-557 <TSU>  
 A:Cross-references: GB:S36200; NID:g249386; PIDN:AA22175.1; PID:g249387  
 A:Experimental source: liver  
 C:Function:  
 A:Description: saposin bind sphingolipids, form hydrophilic complexes and make them acc  
 A:Pathway: sphingolipid catabolism  
 A>Note: saposin A and C (SAP-2) activate hydrolysis of glucocerebroside by beta-glucosy  
 A>Note: saposin B (SAP-1) activates hydrolysis of galactocerebroside sulfate by arylsulfa  
 C:Superfamily: saposin; saposin repeat homology  
 C:Keywords: alternative splicing; glycoprotein; lysosomal storage disease; lysosome; sph  
 F:1-16/Domain: signal sequence #status predicted <SIG>  
 F:1-557/Product: prosaposin #status predicted <PRO>  
 F:55-148/Domain: saposin repeat homology <SAP1>  
 F:55-143/Product: saposin repeat homology <SAP1>  
 F:189-283/Domain: saposin A #status predicted <SAP1>  
 F:184-276/Product: saposin repeat homology <SAP2>  
 F:309-400/Domain: saposin B #status predicted <SAP1>  
 F:313-392/Product: saposin repeat homology <SAP3>  
 F:444-525/Domain: saposin C #status predicted <SAP4>  
 F:440-517/Product: saposin D #status predicted <SAPD>  
 F:63-138.66-132.94-108.197-273,200-267,229-240,317-390,320-384,348-359,442-515,445-509,4  
 F:60,214,334,379,459/Binding site: carbohydrate (asn) (covalent) #status predicted

Query Match 69.4%; Score 1915.5; DB 1; Length 557;  
 Best Local Similarity 63.5%; Pred. No. 1.2e-118;  
 Matches 350; Conservative 77; Mismatches 89; Indels 35; Gaps 3;

2 ASLIGALAPVIGLKECTRGSAVNCQVKTASDCGAVHCLQTVNKPVTSLPCDICK 61  
 8 ASLALATLTPVQDPTCTGSGSAVLCDVKTAVDCGAVHCLQTVNKPVTSLPCDICK 67  
 62 DVTTAAGMLKONATEEELIYLEKTCMDLPRKPMASCKEIVDSYLVILLDIKESNR 121  
 68 TVVTEAGNLLKONATEEELIYLEKTCMDLPRKPMASCKEIVDSYLVILLDIKESNR 127  
 122 PEVGSALNLCSLQNLHAEINHQKLESNKIPBELDTEVVAFFMANIPLLIYPODGRS 181  
 128 PEVGSALNLCSLQNLHAEINHQKLESNKIPBELDTEVVAFFMANIPLLIYPODGRS 186  
 182 KPQPKDNGVCDQCIOMVTDIQTAVRINSTPQALVEHVEKCDRLGFGMADICKNYISQ 241  
 187 QPQPKANEVCDQCKMLVSDVQVAVKTNSSFTQGVDFHVKEDCDRLGFGVSDICKNYVDQ 246  
 242 YSEIATQMMGHN--QPREICALVGFCDVEKEMPQCTIVPAKVASKVIVIPALIEVEPIK 298  
 247 YSEVCCVIMHMVQDQPKKEICVLAFCNEVRAVPMKTLIVPATETIKIILPALIEVMDPEQ 306  
 299 HEVPAKSVYCEVCEELVKEVTKLIDNNKTEKELIDAFDKKCSKLPKSLSEECQEVDTY 358  
 307 NLVQAHNVITLCQTCQFVNNKSEELIVNNAATELVKGLSNACGLPDPARTKCEVVGTF 366  
 359 GSSILSILIEVSPELVCSMLHLCG-----TRL 387  
 367 GPSLIDIFIHVNPSSLCGVIGLCARPELVLEQAPALVSLKLEPPPPQAPQKQ 426  
 388 PALTVHTQPKDGFCEVCKGLVGLDRLNTEKSTKQELIAALEKGSFLPDPYQKQCDQ 447  
 427 SALPAHVPPQKNGFCEVCKGLVGLDRLNTEKSTKQELIAALEKGSFLPDPYQKQCDQ 466  
 448 FVAEYEVLLIEIVVNDPSFVCLKIGACPSAHKPLIGTEKCIWGSRYVQNTETAAQCN 507  
 487 FVAEYEVLLIEIVVNDPSFVCLKIGACPSAHKPLIGTEKCIWGSRYVQNTETAAQCN 546  
 508 AVEHCKRHVNN 518  
 547 AVDHCKRHVNN 557

P109 protein - silkworm  
 C:Species: Bombyx mori (silkworm)  
 C:Update: 22-Jan-1999 #sequence\_revision 22-Jan-1999 #text\_change 21-Jul-2000  
 C:Accession: T00207  
 R:Tambunan, J.; Chang, P.K.; Li, H.; Natori, M.  
 Gene 212, 287-293, 1998  
 A:Title: Molecular cloning of a cDNA encoding a silkworm protein which contains the cons  
 A:Reference number: Z14124; MUID:98278844; PMID:9611271  
 A:Accession: T00207  
 A:Status: Preliminary; translated from GB/EMBL/DBJ  
 A:Molecule type: mRNA  
 A:Residues: 1-965 <TAN>  
 A:Cross-references: EMBL:AB008449; NID:g2575864; PIDN:BA23126.1; PID:g2575865  
 C:Superfamily: saposin repeat homology  
 F:778-870/Domain: saposin repeat homology <SAP3>

Query Match 18.5%; Score 511; DB 2; Length 965;  
 Best Local Similarity 24.6%; Pred. No. 7.9e-26;  
 Matches 139; Conservative 104; Mismatches 229; Indels 94; Gaps 18;

13 VLGLKECTRGSAVNCQVKTASDCGAVHCLQTVNKPVTSLPCDICKVTTAGD 69  
 170 LKGSRTWGPSTWCSNFTGRCNATPHCINRWKMTFPEDDNDCQICLDMVQKARD 229  
 70 MLKONATEEELIYLEKTCMDLPRKPMASCKEIVDSYLVILLDIKESNRPGVCSAL 129  
 230 QLOSNETODIKKEVFGSCGLIPKIFVAGCKMLDEFVELIETLASNN-POAVCSVA 288  
 130 NLCS--LQKHAEINHQKLES-----NKIPBELDTEVVAFFMA----- 167  
 289 GLCNNAKIDPLVDVNAQRELAGCYNCCQTVGVARKKEDETKEDFLVGLLQVGRNMS 348  
 168 ---NIPLLYP-----QDQ-----PRSKPQKNGD 190  
 349 LSDSCMLIFKYENTLEAVKKNLDPBGICHVSQGSYKFNHDEFTPEQMVQYSATD 408  
 191 V--QODICQVTDIQVAVRINSTPQALVEHVEKCDRLGFGMADICKNYISQSEIATQ 249  
 409 VPCEFEQVLKRLVDLVANTTELE-FYKVLQGLCKQKQK-FKDECHLAAQYVYVINF 466  
 250 MM-HMQKEICALVGFCDVEKEMPQCTIVPAKVASKVIVIPALIEVEPIK 288  
 467 LVSDLPKPAFTCGMIGICGNLTSAPISPLVARELVKVPKILGAEBSKIAVPLAKQMP 526  
 289 ALEIVE--PIKHEVPA-KSDYVCEVCEFLVKEVTKLIDNNKTEKELIDAFDKKCSKLPK 345  
 527 ASAAVSVPLEEMFVAAPQSKAACAFQCFILHYLQVQSLDRTEDKVKAAVQEAQDALPD 586  
 346 SLSECCQEVDPYGSILSILIEVSPELVCSMLHLCGTRLPALVHTVQPKDGFCEV 405  
 587 ALNGECKEFVQYGSAAVLAIVQELIDPASVCPALQICQTE-EIRRVAVSEKSN--CPL 643  
 406 CCKLVGYIDRLNENKSTKQELIAALEKGSFLPDPYQKQCDQFVAEYEVLLIEIVVND 465  
 644 CLFAVEQLAESVILKNRSEINERKALDGLCTRLSQKQSECIDFDTYSQGLVNTLVADNN 703  
 466 PSFVCLKIGAC-PSAHKPLIGTEKCI 490  
 704 AKETIVFLKCRDQHLDELKLTSSSI 729

RESULT 5  
 A49475  
 Cerebroside sulfatase activator - pig  
 C:Species: Sus scrofa domestica (domestic pig)  
 C:Date: 24-Feb-1994 #sequence\_revision 18-Nov-1994 #text\_change 17-May-1996  
 C:Accession: A49475  
 R:Stevens, R.L.; Fauli, K.F.; Conklin, K.A.; Green, B.N.; Fluharty, A.L.  
 Biochemistry 32, 4051-4059, 1993  
 A:Title: Porcine cerebroside sulfatase activator: further structural characterization and  
 A:Reference number: A49475; MUID:93229506; PMID:84771613  
 A:Accession: A49475  
 A:Status: Preliminary

A.Molecule type: protein  
A.Residues: 1-79 <STB>  
A.Experimental source: kidney  
A.Note: sequence extracted from NCBI backbone (NCBI:P129597)  
C.Superfamily: saposin repeat homology  
F.1-79/Domain: saposin repeat homology <SAP>

Query Match 14.1% Score 389; DB 2; Length 79;  
Best Local Similarity 88.6%; Pred. No. 3.8e-19;  
Matches 70; Conservative 5; Mismatches 4; Indels 0; Gaps 0;

QY 189 GDVCCDCIQMTYDTIQAVRTNSTFVQALVEHYKECDRLGPGMADICKNYISQYSEIAIQ 248  
DB 1 GDVCCDCIQMTYDTIQAVRTNSTFVQALVEHYKECDRLGPGMADICKNYISQYSEIAIQ 60  
QY 249 MMMHMQPKDICKLVGFCDE 267  
DB 61 MMMHMQPKDICKLVGFCDE 79

## RESULT 6

S02766  
pulmonary surfactant protein B precursor - rat  
C.Species: Rattus norvegicus (Norway rat)  
C.Date: 28-Feb-1990 #sequence\_revision 28-Feb-1990 #text\_change 20-Aug-1999  
C.Accession: S02766  
R.Emile, P.A.; Shannon, J.M.; Mason, R.J.; Fisher, J.H.  
Biochim. Biophys. Acta 994, 215-221, 1989  
A.Title: cDNA and deduced amino acid sequence for the rat hydrophobic pulmonary surfactant  
A.Reference number: S02766; MUID:89150284; PMID:2920185  
A.Accession: S02766  
A.Molecule type: mRNA  
A.Residues: 1-376 <EMR>  
A.Cross-references: EMBL:X14778; NID:957284; PIDN:CAA32885.1; PID:957285  
C.Superfamily: pulmonary surfactant protein B; saposin repeat homology  
F.1-19/Domain: signal sequence #status predicted <SIG>  
F.20-190/Domain: propeptide #status predicted <PRO>  
F.59-151/Domain: saposin repeat homology <SAP1>  
F.190-277/Domain: saposin repeat homology <SAP2>  
F.191-269/Product: pulmonary surfactant protein B #status predicted <MAT>  
F.286-371/Domain: saposin repeat homology <SAP3>

Query Match 13.5% Score 372; DB 2; Length 376;  
Best Local Similarity 22.7%; Pred. No. 3.5e-17;  
Matches 115; Conservative 74; Mismatches 151; Indels 166; Gaps 17;

QY 5 LGALAGPYLGIKECTRGSAVWCQNVKTAASDCGAVGHCLOTWNKPTVKSLPDCICKDV 64  
DB 18 LGALAGPYLGIKECTRGSAVWCQNVKTAASDCGAVGHCLOTWNKPTVKSLPDCICKDV 74  
QY 65 TRAGMLKDNATEEELVLEKTCWMLPEKPNMSASCKEIVDSYLPVLIIDIKGMSRPG 124  
DB 75 HLLTMTREDAQODTIRKFLQECEDILPLKLLVPRRCQVLDVLLVDFPGOI-KPRA 133  
QY 125 VCSALNTEESLGKHLAELNHQKLSNKRLPELDMTEVAVPFMANPLLLLPDGRSKXP 184  
DB 134 ICSHWGLC-----PL--GQTFEQKPE 153  
QY 185 PKDNGDVCCDCIQMTYDTIQAVRTNSTFVQALVEHYKECDRLGPGMADICKNYISQYSE 244  
DB 154 ----- 153  
QY 245 IAIQMMHMQPKDICKLVGFCDEVKEMQTLVPAKVASKNVPALE--LVEP----- 295  
DB 154 -----MLDIAIPPLANKVLPALPGAFLLRPPPTOD 185  
QY 296 IKGEVPAKSDVYCEVCEFLVKEVTKLIDNNKTEKIID-APDKMSKLPKSLSECCOV 354  
DB 186 LSEQQLPIPLP-FCMLCRTLLIKRQAVI-----PKGVLAVAVSVCCHVPLVVGICCOL 239  
QY 355 VDTYSSILSLLEEVSPELVCSMLHCS-----GTRLPALTVAHTQ-PKDGFCFVCKK 408  
DB 240 AERYTVLLLDLALGIVVPLQVGVLRGTADAIGPALPALPLLEKMPLODTTECHFKCS 299

QY 409 LVGYLDRRLKRNSTKQIILALEKGC-SFLDPYXOCDFVAXEYPLIILVEWMPDS 467  
DB 300 VI-----NQAMNTSQAMPQAMQACLRFWLD--RQKCEQVEQMPOLLALVPSQDAH 352  
QY 468 FVCEKIGACPSAHKRLGTEKCIWGP 493  
DB 353 TSCQALGVCEAPASPL-----QCFQTP 374

## RESULT 7

S21770  
saposin-C - bovine  
C.Species: Bos primigenius taurus (cattle)  
C.Date: 19-Mar-1997 #sequence\_revision 19-Mar-1997 #text\_change 31-Oct-1997  
C.Accession: S21770  
R.Sano, A.; Mizuno, T.; Kondoh, K.; Hinenno, T.; Ueno, S.; Kakimoto, Y.; Morita, N.  
Biochim. Biophys. Acta 1120, 75-80, 1992  
A.Title: Saposin-C from bovine spleen; complete amino acid sequence and relation between  
A.Reference number: S21770; MUID:92207994; PMID:1554743  
A.Accession: S21770  
A.Status: preliminary  
A.Molecule type: protein  
A.Residues: 1-80 <SAN>  
C.Superfamily: saposin; saposin repeat homology

Query Match 12.4% Score 343; DB 2; Length 80;  
Best Local Similarity 79.7%; Pred. No. 4.1e-16;  
Matches 63; Conservative 10; Mismatches 6; Indels 0; Gaps 0;

QY 305 SDVCEVCEFLVKEVTKLIDNNKTEKIIDLPDKMSKLPKSLSECCOVDTYSSILIS 364  
DB 1 ADIYQVEEFVYKVAKIDNNKTEEELHLDKVCSTLPSLABQCEVDTYRSILS 60  
QY 365 ILLEEVSPELVCSMLHCS 383  
DB 61 ILLEEVSPELVCSMLHCS 79

## RESULT 8

LNHUB  
pulmonary surfactant protein B precursor [validated] - human  
N.Alternate names: pulmonary surfactant proteolipid SP-B; pulmonary surfactant-associated  
C.Species: Homo sapiens (man)  
C.Date: 31-Dec-1990 #sequence\_revision 31-Dec-1990 #text\_change 08-Dec-2000  
C.Accession: A31361; A28461; A27794; A27592; J00162; S21134  
R.Pilot-Matias, T.J.; Kleier, S.E.; Fox, J.L.; Kropp, K.; Glaeser, S.W.; Whitsett, J.A.  
DNA 8, 75-86, 1989  
A.Title: Structure and organization of the gene encoding human pulmonary surfactant protein  
A.Reference number: A31361; MUID:89170128; PMID:2924687  
A.Accession: A31361  
A.Molecule type: DNA  
A.Residues: 1-381 <PIU>  
A.Cross-references: GB:M24461  
A.Note: the codon given for residue 131 (ATT) is inconsistent with the authors' translation  
R.Jacobs, K.A.; Phelps, D.S.; Steinhilber, R.; Fisch, J.; Kriz, R.; Miltsock, L.; Dougherty, J.  
J. Biol. Chem. 262, 9808-9811, 1987  
A.Title: Isolation of a cDNA clone encoding a high molecular weight precursor to a 6-kDa  
A.Reference number: A28461; MUID:87250653; PMID:3597440  
A.Accession: A28461

A.Molecule type: mRNA  
A.Residues: 1-227, A, 229-381 <JAC>  
A.Cross-references: GB:J02761; NID:g190673; PIDN:AAA60212.1; PID:g190674  
A.Note: part of this sequence, including the amino end of the mature protein, was confirmed  
R.Glaeser, S.W.; Korfhagen, T.R.; Weaver, T.; Pilot-Matias, T.; Fox, J.L.; Whitsett, J.A.  
Proc. Natl. Acad. Sci. U.S.A. 84, 4007-4011, 1987  
A.Title: cDNA and deduced amino acid sequence of human pulmonary surfactant-associated protein  
A.Reference number: A27794; MUID:87231940; PMID:3035561  
A.Accession: A27794  
A.Molecule type: mRNA  
A.Residues: 'EPR', 99-117, 'L', 319-381 <GLA>  
A.Cross-references: GB:M16764; NID:g338410; PIDN:AAA88099.1; PID:g338411

A>Note: 131-Ile was also found  
A>Note: Part of this sequence, including the amino end of the mature protein, was confix  
R.Revak, S.D.; Merritt, T.A.; Degryse, E.; Stefani, L.; Courtney, M.; Hallman, M.; Coch  
J. Clin. Invest. 81, 826-833, 1988  
A>Title: Use of human surfactant low molecular weight apoproteins in the reconstitution  
A:Reference number: A27592; MUID:88139786; PMID:3343343  
A:Accession: A27592  
A:Molecule type: mRNA  
A:Residues: 139-177, 'V', 179-227, 'A', 228-381 <REV>  
A:Cross-references: GB:M19097  
A>Note: Part of this sequence, including the amino end of the mature protein, was confix  
R.Mizumoto, M.; Adachi, H.  
Sapporo Igaku Zasshi 56, 731-742, 1987  
A>Title: Primary structure of a hydrophobic 6kDa apoprotein (SP6) of human pulmonary su  
A:Reference number: J00162  
A:Accession: J00162  
A:Molecule type: protein  
A:Residues: 201-207, 'X', 209-210, 'X', 212-227, 'A', 229-334, 'X', 236-245, 'X', 247, 'L', 249-253,  
R.Johansson, U.; Joernvall, H.; Curstedt, T.  
FEBS Lett. 301, 165-167, 1992  
A>Title: Human surfactant polypeptide SP-B. Disulfide bridges, C-terminal end, and pep  
A:Reference number: S21134; MUID:92233937; PMID:1568474  
A:Accession: S21134  
A>Status: preliminary  
A:Molecule type: protein  
A:Residues: 201-227, 'T', 229-279 <JOH>  
A>Note: 228-Ala was also found  
C/Comment: Pulmonary surfactant is a complex of phospholipids and proteins that lowers t  
A:Gene: GDB:SFTPB; SFTPB; SP-B  
A:Cross-references: GDB:120374; OMIM:178640  
A:Map position: 2p12-2p11.2  
A:Intons: 22/1; 65/3; 89/3; 131/3; 194/3; 224/3; 286/1; 334/3; 361/3  
C/Superfamily: pulmonary surfactant protein B; saposin repeat homology  
C/Keywords: alveolar proteinosis; gaseous exchange; glycoprotein; lipoprotein; lung; pul  
F:1-18/Domain: signal sequence #status predicted <SIG>  
F:19-200/Domain: propeptide #status predicted <PRO>  
F:61-153/Domain: saposin repeat homology <SAP1>  
F:200-287/Domain: saposin repeat homology <SAP2>  
F:201-279/Domain: pulmonary surfactant protein B, 9K form #status predicted <SP9>  
F:201-256/Domain: pulmonary surfactant protein B, 6K form #status experimental <SP6>  
F:291-376/Domain: saposin repeat homology <SAP3>  
F:69-143, 72-137, 100-112, 299-366, 302-360, 335-335/Disulfide bonds: #status predicted  
F:129, 311/Binding site: carbohydrate (asn) (covalent) #status predicted  
F:208-277, 211-271, 235-246/Disulfide bonds: #status experimental  
F:248/Disulfide bonds: interchain #status experimental

Query Match 12.3%; Score 339.5; DB 1; Length 381;  
Best Local Similarity 21.4%; Pred. No. 4, 9e-15;  
Matches 109; Conservative 71; Mismatches 169; Indels 161; Gaps 16;

QY 1 LASLIGALAPVGLGKCTGSAVWCQNVKTASDCGAVKICLQTVNKKPTVSLPDCIC 60  
DB 14 LPTLGGPETAAMWTTSSLAACQGFPCWCSLEQALQCRALGHCLQEVGVHGAIDL-CQEC 72  
QY 61 KDVTAAQDMKDNATEBEILVLEKTCDMLEPKNMSASKEIVDSYLPVILDIKGMRS 120  
DB 73 EDIHIILNKAKKAIPODTMKKFLQECNVLPKLPLPQCQNVLDYFPIVIDYFQVQTD 132  
QY 121 RRGVCSALNLCESLQKHLAELNHQKLESNKIPELDMTEVVAFFMANIPLLVLPQDGR 180  
DB 133 SNG-ICMHLGLCKSRQ----- 147  
QY 181 SKPOKNDGVCDQCIQWVDIQTAVRTNSTFQVALVEHKECDRLGPGMADICKNYIS 240  
DB 148 --PEPEDE-----PQMSD----- 158  
QY 241 QYSEIALQMMHMQKEICALVGFCDVEKEMPMQTLVPAKVASKNYIPALEVEPIKKE 300  
DB 159 -----PLPK-----PLRDLPLDLDKLVLPVLP--GALQARP 189  
QY 301 VPAKSDV-----YCEVCEFLVEVTKLIDNKKTEKEILDADFKNCKSLPKSLSEE 350

DB 190 GPHOTDLEEQGFPIPLPFCMLCRALLIRIQAMIRK-----ALRVAQVCRVPLVAGI 245  
QY 351 COEVDYTGSSILSLILEVSPELVCSMLHICS-----GTRLPALTVAHTQPKDGFCEV 405  
DB 246 CQCLAEKRSVILLDTLGRMPLPOLVCRVLRCSWDSAGPRSP---TGEWLPDRSE-CHL 301  
QY 406 CKLVGYLDRLNLEKSTKQEFILALEKGC--SFLLPDYQKOCQFVAVEFVLIILEV 463  
DB 302 CMSTV-----TQANSSQEQALPQMLQACGSMWD---KKCKQFVEQHTPQLTLVPRG 353  
QY 464 MDPSPVCLTKGACPSAHKPLIGTEKTIWGP 493  
DB 354 WDAHTTQCALGVCGTGMSSPL-----QCIHSP 379

RESULT 9  
A29072  
pulmonary surfactant protein SP 18 precursor - dog (fragment)  
C/Species: Canis lupus familiaris (dog)  
C/Date: 31-Mar-1989 #sequence\_revision 31-Mar-1989 #text\_change 20-Aug-1999  
C/Accession: B29072; A29072  
R:Hawgood, S.; Benson, B.J.; Schilling, J.; Damm, D.; Clements, J.A.; White, R.T.  
Proc. Natl. Acad. Sci. U.S.A. 84, 66-70, 1987  
A>Title: Nucleotide and amino acid sequences of pulmonary surfactant protein SP 18 and ev  
A:Reference number: A29072; MUID:87092398; PMID:3467361  
A:Accession: B29072  
A:Molecule type: mRNA  
A:Residues: 1-363 <HAM>  
A:Cross-references: GB:M15170; NID:g164077; PIDN:AAA0893.1; PID:g164078  
A:Accession: A29072  
A:Molecule type: protein  
A:Residues: 182-210 <HA2>  
C/Superfamily: pulmonary surfactant protein B; saposin repeat homology  
F:1-14/Domain: signal sequence #status predicted <SIG>  
F:15-180/Domain: propeptide #status predicted <PRO>  
F:54-146/Domain: saposin repeat homology <SAP1>  
F:180-267/Domain: saposin repeat homology <SAP2>  
F:181-363/Domain: pulmonary surfactant protein SP 18 #status experimental  
F:273-358/Domain: saposin repeat homology <SAP3>

Query Match 11.1%; Score 305; DB 2; Length 363;  
Best Local Similarity 21.6%; Pred. No. 8, 7e-13;  
Matches 108; Conservative 75; Mismatches 159; Indels 158; Gaps 19;

QY 5 LGAALAGVGLGKCTGSAVWCQNVKTASDCGAVKICLQTVNKKPTVSLPDCICDV 63  
DB 13 LGAADMSAPSLA---CARGPAFWCSLEQALQCRALGHCLQEVGVHGAIDL-CQECODI 68  
QY 64 VTAAGDMKDNATEBEILVLEKTCDMLEPKNMSASKEIVDSYLPVILDIKGMRSRG 123  
DB 69 VAILTKMKKAIPODMVAKKFLHEKCDVLPKLTLPPQCHMIGTFPVVVDYFQSQIN-PK 127  
QY 124 EYCSALNLCESLQKHLAELNHQKLESNKIPELDMTEVVAFFMANIPLLVLPQDGRSKP 183  
DB 128 IICKHGLC---KGLPEPEGESELSPLDLKTLPEL----- 162  
QY 184 QPKDNDGVCDQCIQWVDIQTAVRTNSTFQVALVEHKECDRLGPGMADICKNYISQYS 243  
DB 163 ----- 177  
QY 244 EIALQMMHMQKEICALVGFCDVEKEMPMQTLVPAKVASKNYIPALEVEPIKKEHVPA 303  
DB 178 E-----QQLP-----LP----- 186  
QY 304 KSDVYCEVCEFLVEVTKLIDNKKTEKEILD-AFDKCKSLPKSLSEECQEVVDYTGSSI 362  
DB 187 ---YCMICRTLLIRIQAMT-----PKGVLAIVTVGQCHVPLVVGIGICQCLAEKRYTLL 237  
QY 363 LSLILEVSPELVCSMLHICS-----GTRLPALTVAHTQPKDGFCEVCKLVGYLDRL 417  
DB 238 LDALLGRMLPOLVGVGLVLRCSHSDSAGPALASPSE-WSPQSKC-QCLCMFVLT-----TQ 290

Qy 418 EKNSKQEIILAEKGC--SFLPDPYOKCDQFAVEYEVILIEVWMDSEFVCLTKIGA 475  
Db 291 AGNHSQKTPQAIIRACLSWMD---RQKCBQFVQGHRIQTLASGGRDATTQALGA 347  
Qy 476 CP5AHKPKLGTREKCIWGPSY 495  
Db 348 CRTFSPL---QCIIHPHF 363

## RESULT 10

146531  
surfactant protein B - rabbit  
C:Species: Oryctolagus cuniculus (domestic rabbit)  
C>Date: 14-Feb-1997 #sequence\_revision 14-Feb-1997 #text\_change 20-Aug-1999  
C:Accession: 146531  
R:Margana, R.K.; Boggaram, V.  
Am. J. Physiol. 268, L481-L490, 1995  
A:Title: Transcription and mRNA stability regulate developmental and hormonal expression  
A:Reference number: 146531, MUID:95208794, PMID:7900830  
A:Accession: 146531  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: mRNA  
A:Residues: 1-369 <MAN>  
A:Cross-references: EMBL:U17106, NID:9642487, PIDN:AAA67934.1, PID:9642488  
C:Genetics:  
A:Gene: SP-B  
C:Superfamily: pulmonary surfactant protein B; saposin repeat homology  
F:61-153/Domain: saposin repeat homology <SAP1>

Query Match 10.6%; Score 293; DB 2; Length 369;  
Best Local Similarity 21.6%; Pred. No. 5-5e-12;  
Matches 110; Conservative 65; Mismatches 165; Indels 170; Gaps 19;

Qy 4 LIGAAALAPVVLGL-----KECTGSAVWCQNVKTAASDCGAVHCLQTQWKNKPTVSLPDC 58  
Db 12 LLLPPLCGPAGVAVATSPACAGGPEFQCGLQKALGHCLQEWGHHGADDL-CQ 70  
Qy 59 ICKDVVTAAGMLKDNATEEBILVYLEKTCMLPRPNNSASCKEIVDSYLPVILDIIGE 118  
Db 71 ECQDVLNLTQWTKKAIPODTRKFLHEBCDVLPLKLLVPOCHHVLADVFPPLTITYF 130  
Qy 119 MSRPVEVSALNLCESLQKHLAELNHOKLSNKPRLDMTEVAVPFAANIPLILYPQDG 178  
Db 131 INAKA-ICQHLGLCQ-----PQSEPPPLD-----PLPDKLVLPFL--GA 167  
Qy 179 PRSKRPQPKNDQVCDICIQWTDIQTAVRNTSTFQVALVHVKESCDRLGPMADICKY 238  
Db 168 LPAPK-----GPHTDLS--- 180  
Qy 239 ISOVSEIAIQMMHMQPEICALVGCDEVKEMPMQTLVPAKVASKNVIPLALEVPEIKK 298  
Db 181 -----AQRFPPL-----PL-- 190  
Qy 299 HEVPAKSDVCEVCEFLVKEVTKLIDNNKTEKEILD-AFDKMSCLPLPSLSECEQVVD 357  
Db 191 -----CMLCKTLAKRIQAMT-----PKGVLAAMAVAGCHVPLVVGICQRLER 235  
Qy 358 YGSSILSLLEBVSDELVCMLHCS-----GTRLPALTVAHTQ--PKDGFCEVCKLV 410  
Db 236 YTVILLLEVLGHVLPOLVCGHVRCSVDSIGQVPTLEALGEMLPDPPE-CRLCMSVT 294  
Qy 411 GYLDRLNEKNSKQEIILAA-----LEKGSFLPDPYOKCDQFAVEYEVILIEVWMD 465  
Db 295 TQA-NNISEQTRPQAVVHACISSQDK-----QECQFVELATPQLSLSRGMD 343  
Qy 466 PSFVCLKIGACPSAHKPKLGTREKCIWGPSY 495  
Db 344 ARAIQALGACAVATISPL-----QCIGSPHF 369

RESULT 11  
A32026  
glucosylceramide beta-glucosidase activator protein SAP-2 - guinea pig

C:Species: Cavia porcellus (guinea pig)  
C>Date: 07-Jun-1990 #sequence\_revision 07-Jun-1990 #text\_change 17-May-1996  
C:Accession: A32026  
R:Sanu, A.; Radin, N.S.; Johnson, L.L.; Tarr, G.E.  
J. Biol. Chem. 263, 19597-19601, 1988  
A:Title: The activator protein for glucosylceramide beta-glucosidase from guinea pig liv  
A:Reference number: A32026, MUID:89066787, PMID:3198642  
A:Accession: A32026  
A:Molecule type: protein  
A:Residues: 1-81 <SAN>  
C:Superfamily: saposin; saposin repeat homology  
F:1-81/Domain: saposin repeat homology <SAP>

Query Match 9.5%; Score 261; DB 2; Length 81;  
Best Local Similarity 61.8%; Pred. No. 1.1e-10;  
Matches 47; Conservative 18; Mismatches 11; Indels 0; Gaps 0;

Qy 307 VYCEVCEFLVKEVTKLIDNNKTEKEILDAPFMCKSLPKPSISECEQVVDYTGSSILSL 366  
Db 3 VTCACACEVAVKQWELIDNNRTEKIHALLDSVCLALPESVSEVCEVVDYTGDSIVALL 62  
Qy 367 LEEVSPBELVCSMLHLC 382  
Db 63 LQEMSPBELVCSSELGLC 78

## RESULT 12

## INRRB

pulmonary surfactant protein B precursor - rabbit  
N:Alternate names: pulmonary surfactant-associated protein-B  
C:Species: Oryctolagus cuniculus (domestic rabbit)  
C>Date: 31-Dec-1990 #sequence\_revision 31-Dec-1990 #text\_change 22-Jun-1999  
C:Accession: A32421  
R:Xu, J.; Richardson, C.; Ford, C.; Spencer, T.; Li-Juan, Y.; Mackie, G.; Hammond, G.; P  
Biochem. Biophys. Res. Commun. 160, 325-332, 1989  
A:Title: Isolation and characterization of the cDNA for pulmonary surfactant-associated  
A:Reference number: A32421, MUID:89228033, PMID:2469419  
A:Accession: A32421  
A:Molecule type: mRNA  
A:Residues: 1-370 <XU>  
A:Cross-references: GB:M24901, NID:9165707, PIDN:AAA31466.1, PID:9165708  
A>Note: the authors translated the codon CCG for residue 184 as Arg and CAG for residue  
C:Comment: Pulmonary surfactant is a complex of phospholipids and proteins that lowers t  
C:Superfamily: pulmonary surfactant protein B; saposin repeat homology  
C:Keywords: alveolar proteinosis; gaseous exchange; glycoprotein; lipoprotein; lung; pul  
F:1-21/Domain: signal sequence #status predicted <SIG>  
F:22-184/Domain: propeptide #status predicted <PRO>  
F:62-154/Domain: saposin repeat homology <SAP1>  
F:184-271/Domain: saposin repeat homology <SAP2>  
F:185-263/Product: pulmonary surfactant protein B, 9K form #status predicted <SP9>  
F:185-240/Product: pulmonary surfactant protein B, 6K form #status predicted <SP6>  
F:280-365/Domain: saposin repeat homology <SAP3>  
F:300/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 9.1%; Score 250; DB 1; Length 370;  
Best Local Similarity 20.7%; Pred. No. 3.8e-09;  
Matches 108; Conservative 63; Mismatches 162; Indels 190; Gaps 21;

Qy 1 LASLIGAAALAPVVLGL-----KECTGSAVWCQNVKTAASDCGAVHCLQTQWKNKPTVSL 55  
Db 10 LLLPPLCGPAGVAVATSPACAGGPEFQCGLQKALGHCLQEWGHHGADDL 69  
Qy 56 PCDICQDVVTAAGMLKDNATEEBILVYLEKTCMLPRPNNSASCKEIVDSYLPVILDI 115  
Db 70 -CORCODVNLITQWTKKAIPODTRKFLHEBCDVLPLKLLVPOCHHVLADVFPPLTITYF 128  
Qy 116 KGMSPREVEVSALNLCESLQKHLAELNHOKLSNKPRLDMTEVAVPFAANIPLILYP 175  
Db 129 QSOINAKA-ICQHLGLCQ-----PQSEPPPLD-----PLPDKLVLPFL-- 166  
Qy 176 QDGPSPKRPQPKNDQVCDICIQWTDIQTAVRNTSTFQVALVHVKESCDRLGPMADIC 235  
Db 167 -GALPAKRGF----- 175

QY 236 KNYISQYSEIALQMMHMQPKICALVGFCDYKEMQTLVPAKVASKNVIALBELVEP 295  
Db 176 -----HQ-----DLQAQFPPIPLP-----190  
QY 296 IKKHEVPAKSDVYCEVCEFLVKEVTKLIDNNKTEKEIIDA-APDMGSKLPKSLSECOY 354  
Db 191 -----LCMLCRITLLKRIQAMI-----PKGYLAAVAVOVCHVAVFLVVGICQCL 233  
QY 355 VDTGSSILSTILLEEVPBELVCSMLHCS-----GTRLPALTVAHTQ--PKDGFCEVCK 407  
Db 234 AERTVTLLEVLGHVLPQVLVGRSSVDSIQVPTLEALPGEMLPQDDE--CRLCM 292  
QY 408 KLVGYIDRLNLEKSTKOEIIAA-----LEKGCSEFLPDPYQKQDQFAVEEPLIILIVE 462  
Db 293 SVTTQA-RNISEQRRPQAVVHACISQGLD-----QCEQGFVAHAAP-----333  
QY 463 VMDPSFVCLKIGACPSAKPLLGT-----KCIWGPSY 495  
Db 334 AAEPAVQGL---GCP---RNLPGEGRVVATLSPLQICQSPHF 370

## RESULT 13

T46069  
hypothetical protein T18N14.110 - Arabidopsis thaliana  
C:Species: Arabidopsis thaliana (mouse-ear cress)  
C>Date: 04-Feb-2000 #sequence\_revision 04-Feb-2000 #text\_change 04-Feb-2000  
C:Accession: T46069  
R:Deleeny, M.; Berger, C.; Cooke, R.; Grellet, F.; Laurie, M.; Mewes, H.W.; Lemcke, K.;  
submitted to the Protein Sequence Database, December 1999  
A:Reference number: Z23013  
A:Accession: T46069  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-213 <DEL>  
A:Cross-references: EMBL:AL132968  
A:Experimental source: cultivar Columbia; BAC clone T18N14  
C:Genetics:  
A:Map position: 3  
A:introns: 31/1; 80/3; 146/3; 166/3  
A>Note: T18N14.110

Query Match 7.7%; Score 213; DB 2; Length 213;  
Best Local Similarity 29.7%; Pred. No. 5, 1e-07;  
Matches 57; Conservative 35; Mismatches 90; Indels 10; Gaps 6;

QY 286 VIPALELVEPIKHEVPAKSDVYCEVCEFLVKEVTKLIDNNKTEKEIIDA-APDMGSKLPK 345  
Db 18 VSDARSFVDSITSEKVSNEKEDV-CTLCEEVTDALSTYLEKNVTOABIIEDLHRCQGL-R 75  
QY 346 SLSECCQEVDTYGGSSILSTILLEEVPBELVCSMLHCSGTRLPALTVAHTQPKDGFCEV 405  
Db 76 GYSQCCISLVVYV-VPLFQLQESQPHYFCRKNLQG--KVVALVEARQDS---CGV 128  
QY 406 CKLVGYIDRLNLEKSTKOEIIAALEKGCSEFLPDPYQKQDQFAVEEPLIILIVEVD 465  
Db 129 CHRVSSELLIKQDPDQDIDIVELLIKGCKSLKN-YEKCKITLVEEGLVILVNAEEFLV 187  
QY 466 PSFVCLKIGACP 477  
Db 188 KNDVCTILRACP 199

## RESULT 14

T48201  
hypothetical protein T20L15.70 - Arabidopsis thaliana  
C:Species: Arabidopsis thaliana (mouse-ear cress)  
C>Date: 20-Apr-2000 #sequence\_revision 20-Apr-2000 #text\_change 20-Apr-2000  
C:Accession: T48201  
R:Bevan, M.; Peters, S.A.; van Steveren, M.; Dirks, W.; Stiekema, W.; Bancroft, I.; Mew  
submitted to the Protein Sequence Database, March 2000  
A:Reference number: Z24488  
A:Accession: T48201

A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-217 <BEV>  
A:Cross-references: EMBL:AL162351  
A:Experimental source: cultivar Columbia; BAC clone T20L15  
C:Genetics:  
A:Map position: 5  
A:introns: 30/1; 79/3; 146/3; 166/3  
A>Note: T20L15.70

Query Match 7.4%; Score 203.5; DB 2; Length 217;  
Best Local Similarity 27.8%; Pred. No. 2, 2e-06;  
Matches 52; Conservative 40; Mismatches 82; Indels 13; Gaps 7;

QY 292 LVEPIKK-HEVPAKSDVYCEVCEFLVKEVTKLIDNNKTEKEIIDA-APDMGSKLPKSLSE 350  
Db 25 LLEPFESAHND---DNQVCELCDKVTYTLVDYIDQDNQNEVEALHISCSQIP--PLKIQ 79  
QY 351 QCEVDTYGGSSILSTILLEEVPBELVCSMLHCSGTRLPALTVAHTQPKDGFCEVCKLY 410  
Db 80 CLSWVDHY-TQLPFTQVSTIKSDQICKRLNLQAV--TPAFASQVHQ---GNCEACRETV 133  
QY 411 GYLDRLNLEKSTKOEIIAALEKGCSEFLPDPYQKQDQFAVEEPLIILIVEVMDPSFVC 470  
Db 134 SEVYTKLKDPEYTKLIRLLLECKSL--NNYQDKCKWFEYEGPLMTLQKFLERKQVC 192  
QY 471 LKIGACP 477  
Db 193 TILHVCIP 199

## RESULT 15

T15677  
hypothetical protein C28C12.5 - Caenorhabditis elegans  
C:Species: Caenorhabditis elegans  
C>Date: 20-Sep-1999 #sequence\_revision 20-Sep-1999 #text\_change 20-Sep-1999  
C:Accession: T15677  
R:Miller, N.  
submitted to the EMBL Data Library, November 1995  
A:Description: The sequence of C. elegans cosmid C28C12.  
A:Reference number: Z18387  
A:Accession: T15677  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 1-402 <MTL>  
A:Cross-references: EMBL:U40797; NID:g1065916; PID:g1065921; PIDN:AAB37548.1; GSPDB:GN00C  
A:Experimental source: strain Bristol N2; clone C28C12  
C:Genetics:  
A:gene: CESP:C28C12.5  
A:Map position: 4  
A:introns: 29/3; 82/3; 124/3; 151/3; 258/3; 318/1; 372/3; 400/3

Query Match 6.4%; Score 176.5; DB 2; Length 402;  
Best Local Similarity 20.3%; Pred. No. 0.00029;  
Matches 75; Conservative 58; Mismatches 156; Indels 81; Gaps 15;

QY 178 GPRSKPQKNDGVCOQDQICQWTDIQTAVRNSTFVQALVEHVEEDRGRPGMAQICN 237  
Db 16 GAQSAFSP-----CECKSKVQNFIDASKDRKMAQLKVS--LSMLC--VGTSHQSDCK 66  
QY 238 YISQYSEIALQMMHMQPKICALVGFCDYKEMQTLVPAKVASKNVIALBELVEPI 296  
Db 67 TLDKIDFLAYKLAAYLADTSVCSKIQWCGESQFSLARLAMLYLKSEIVA---NDNI 123  
QY 297 KKHVEPAKSDVYCEVCEFLVKEVTKLIDNNKTEKEIIDA-APDMGSKLPKSLSECOEVD 356  
Db 124 MRQEV-----CDECASTAQIGKLVGDEFTYAVASTLQRFVCSAGKAMHAGC-----N 172  
QY 357 TYGSSILSTILLEEV-----SPELVCSMLHCSGTRLPALTVAHTQPKDGFCEVCKLVG 411  
Db 173 IFVSVIPDILMTWKDQFTEKELMGCSNMGICSAITSKPARARAPKQAS---EMKSM-- 226  
QY 412 YLDRLNLEKSTKOEIIAALEKGCSEFLPDPYQKQ----- 444

```

Db      227  ----GMVTSINSEELMSCFE---CTLSADMLLGEFIDKRGSTADDICTVACNKNVAVNTDGO 280
OY      445  CDQFAAEFEPLVILIELIVEMDPSFVCLKTIGACPSAHKPLLG---TEKCTWGPSYWCNTE 501
Db      281  CNDPFAHMTSIVTLFLITYNGFDGRGICTVMHSCCKENALVENAMSEKVMLG---CEN-- 334
OY      502  TAAOCNAVEH 511
Db      335  ----CKAVEH 340

```

Search completed: May 5, 2004, 13:34:35  
Job time : 22.661 secs

**THIS PAGE BLANK (USPTO)**



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 5, 2004, 13:16:48 ; Search time 12.7901 Seconds

(without alignments)  
2108.841 Million cell updates/sec

Title: US-09-743-684a-1\_COPY\_7\_524

Sequence: 1 LASILGALAGPVLGKRECT.....NTEPAQCNAYEHCKRHVWN 518

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 141681 seqs, 52070155 residues

Total number of hits satisfying chosen parameters: 141681

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt\_42:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2759	100.0	524	1	SAP_HUMAN
2	2405.5	87.2	525	1	SAP_BOVIN
3	1974	71.5	554	1	SAP_RAT
4	1916.5	69.5	557	1	SAP_MOUSE
5	1676.5	60.8	518	1	SAP_CHICK
6	393	14.2	80	1	SAP_PIG
7	378	13.7	377	1	PSPB_MOUSE
8	372	13.5	376	1	PSPB_RAT
9	340.5	12.3	381	1	PSPB_HUMAN
10	305	11.1	363	1	PSPB_CANPA
11	290	10.5	370	1	PSPB_RABIT
12	261	9.5	81	1	SAP_CAVPO
13	173.5	6.3	513	1	ASPR_CUCPE
14	158.5	5.7	508	1	ASPR_HORVU
15	154.5	5.6	473	1	CYPL_CYNCA
16	154	5.6	496	1	ASPR_ORYSA
17	143.5	5.2	1175	1	GLG1_MOUSE
18	142.5	5.2	509	1	APR1_ORYSA
19	138.5	5.0	1171	1	GLG1_RAT
20	137.5	5.0	1179	1	GLG1_HUMAN
21	137.5	5.0	1160	1	GLG1_CRIGR
22	127.5	4.6	1142	1	GLG1_CHICK
23	120	4.3	578	1	EZRA_ENTRA
24	119	4.3	975	1	SECB_MOUSE
25	118	4.3	857	1	12C1_HUMAN
26	117	4.2	857	1	12C1_MOUSE
27	116.5	4.2	1216	1	PIB1_MOUSE
28	116	4.2	975	1	SECB_RAT
29	114.5	4.2	971	1	RBECK_HUMAN
30	114	4.1	984	1	NOR_DROME
31	113.5	4.1	974	1	SECB_HUMAN
32	113	4.1	418	1	TEKI_HUMAN
33	113	4.1	1102	1	MYSC_CHICK

34	113	4.1	2869	1	RBP1_PLAUV	Q00798 plasmodium
35	112	4.1	1184	1	PBI2_HUMAN	P98098 homo sapien
36	112	4.1	1216	1	PBI1_BOVIN	P10894 bos taurus
37	112	4.1	1216	1	PBI1_RAT	P10687 rattus norv
38	111.5	4.0	8545	1	ANCI1_CAEEL	Q9h4m4 caenorhabdi
39	110.5	4.0	3911	1	AKA9_HUMAN	Q99996 h-a-kinase
40	110	4.0	476	1	VTDB_RABIT	P53789 oryctolagus
41	110	4.0	1216	1	PBI1_HUMAN	Q9nq66 homo sapien
42	109	4.0	861	1	12C4_HUMAN	Q9nck5 homo sapien
43	109	4.0	1557	1	DVAL_DICTY	Q24702 dictyocaulu
44	108.5	3.9	3674	1	SPCR_HUMAN	Q9nrc6 homo sapien
45	108.5	3.9	3678	1	DMD_MOUSE	P11531 mus musculu

## ALIGNMENTS

RESULT 1  
SAP\_HUMAN STANDARD: PRT: 524 AA  
AC P07602; P07292; P15793; P78538; P78541; P78546; P78547; P78558;  
AC Q92739; Q92740; Q92741; Q92742;  
DT 01-APR-1988 (Rel. 07, Created)  
DT 01-APR-1990 (Rel. 14, Last sequence update)  
DT 10-OCT-2003 (Rel. 42, Last annotation update)  
DE Proactivator polypeptide precursor [contains: Saposin A (Protein A);  
DE Saposin B (Sphingolipid activator protein 1) (SAP-1) (Cerebroside  
DE sulfate activator) (CSact) (Dispersin) (Sulfatide/GM1 activator);  
DE Saposin C (Co-beta-glucosidase) (AI activator) (Glucosylceramidase  
DE activator) (Sphingolipid activator protein 2) (SAP-2); Saposin D  
DE (Protein C) (Component C)].  
GN PSAP.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP TISSUE=Liver;  
RC MEDLINE=90129043; PubMed=2515150;  
RA Roman E.G., Grabowski G.A.;  
RT "Molecular cloning of a human co-beta-glucosidase cDNA: evidence that  
RT four sphingolipid hydrolase activator proteins are encoded by single  
RT genes in humans and rats.";  
RL Genomics 5:486-492(1989).  
[2]  
RP TISSUE=Brain, Eye, and Skin;  
RC MEDLINE=89255151; PubMed=2498298;  
RA Nakano T., Sandhoff K., Stuenkel J., Christomanou H., Suzuki K.;  
RT "Structure of full-length cDNA coding for sulfatide activator, a  
RT co-beta-glucosidase and two other homologous proteins: two alternate  
RT forms of the sulfatide activator.";  
RL J. Biochem. 105:152-154(1989).  
[3]  
RP TISSUE=Brain, Eye, and Skin;  
RC MEDLINE=22388257; PubMed=12477932;  
RA Strussberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,  
RA Klausner R.D., Collins F.S., Wagner L., Shemmen C.M., Schuler G.D.,  
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,  
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,  
RA Diatchenko L., Marnett K., Farmer A.A., Rubin G.M., Hong L.,  
RA Stapleton M., Soares W.B., Bonaldo M.P., Casavant T.L., Schetz T.E.,  
RA Bawa S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,  
RA Boeck S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,  
RA Villalón D.K., Morley K.C., Hale S., Garcia A.M., Gay L.J., Hulik S.W.,  
RA Faney J., Helton E., Ketteman M., Madan A., Rodríguez S., Sanchez A.,  
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,  
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,  
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,  
RA Butterfield V.S.N., Krzywinski M.I., Skalske U., Smallos D.E.,

- RA Schmerch A., Schern J.E., Jones S.J.M., Marra M.A.;  
RT "Generation and initial analysis of more than 15,000 full-length  
RT human and mouse cDNA sequences.";  
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
- RA [4]  
RC SEQUENCE OF 59-125 AND 304-513 FROM N.A.  
RC TISSUE=Brain;  
RX MEDLINE=91192146; PubMed=2013321;  
RA Holtschmidt H., Sandhoff K., Fuert W., Kwon H.Y., Schnabel D.,  
RA Suzuki K.;  
RT "The organization of the gene for the human cerebroside sulfate  
RT activator protein.";  
RL FEBS Lett. 280:267-270(1991).
- RA [5]  
RC SEQUENCE OF 164-524 FROM N.A.  
RX MEDLINE=8068647; PubMed=2825202;  
RA Dewi N.N., Wenger D.A., O'Brien J.S.;  
RT "Nucleotide sequence of cloned cDNA for human sphingolipid activator  
RT protein 1 precursor.";  
RL Proc. Natl. Acad. Sci. U.S.A. 84:8652-8656(1987).
- RA [6]  
RC PARTIAL SEQUENCE OF 60-142.  
RX MEDLINE=89240739; PubMed=2717620;  
RA Morimoto S., Martin B.M., Yamamoto Y., Kretz K.A., O'Brien J.S.,  
RA Kishimoto Y.;  
RT "Saposin A: second cerebroside activator protein.";  
RL Proc. Natl. Acad. Sci. U.S.A. 86:3389-3393(1989).
- RA [7]  
RC SEQUENCE OF 195-263 FROM N.A.  
RX MEDLINE=86130593; PubMed=2668718;  
RA Dewi N.N., Wenger D.A., Fujibayashi S., Donoviel M., Esch F.,  
RA Hill F., O'Brien J.S.;  
RT "Molecular cloning of the sphingolipid activator protein-1 (SAP-1),  
RT the sulfatide glucosylase activator.";  
RL Biochem. Biophys. Res. Commun. 134:989-994(1986).
- RA [8]  
RC SEQUENCE OF 195-274.  
RC TISSUE=Kidney;  
RX MEDLINE=91006165; PubMed=2209618;  
RA Furst W., Schubert J., Machleidt W., Meyer H.E., Sandhoff K.;  
RT "The complete amino-acid sequences of human ganglioside GM2 activator  
RT protein and cerebroside sulfate activator protein.";  
RL Eur. J. Biochem. 192:709-714(1990).
- RA [9]  
RC SEQUENCE OF 195-274.  
RX MEDLINE=89207118; PubMed=3242555;  
RA Kleinschmidt T., Christomanou H., Braunitzer G.;  
RT "Complete amino-acid sequence of the naturally occurring A2 activator  
RT protein for enzymic sphingomyelin degradation: identity to the  
RT sulfatide activator protein (SAP-1).";  
RL Biol. Chem. Hoppe-Seyler 369:1361-1365(1988).
- RA [10]  
RC SEQUENCE OF 311-390.  
RX MEDLINE=88163077; PubMed=3442600;  
RA Kleinschmidt T., Christomanou H., Braunitzer G.;  
RT "Complete amino-acid sequence and carbohydrate content of the  
RT naturally occurring glucosylceramide activator protein (A1 activator)  
RT absent from a new human Gaucher disease variant.";  
RL Biol. Chem. Hoppe-Seyler 368:1571-1578(1987).
- RA [11]  
RC SEQUENCE OF 407-484.  
RX MEDLINE=8900190; PubMed=3048308;  
RA Furst W., Machleidt W., Sandhoff K.;  
RT "The precursor of sulfatide activator protein is processed to three  
RT different proteins.";  
RL Biol. Chem. Hoppe-Seyler 369:317-328(1988).
- RA [12]  
RC PARTIAL SEQUENCE OF 405-484.  
RX MEDLINE=89025876; PubMed=2845979;  
RA Morimoto S., Martin B.M., Kishimoto Y., O'Brien J.S.;  
RT "Saposin D: a sphingomyelinase activator.";  
RL Biochem. Biophys. Res. Commun. 156:403-410(1988).
- RA [13]
- RP SEQUENCE OF 17-26.  
RC TISSUE=Milk;  
RX MEDLINE=92068206; PubMed=1958198;  
RA Kondoh K., Hineno T., Sano A., Kakimoto Y.;  
RT "Isolation and characterization of prosaposin from human milk.";  
RL Biochem. Biophys. Res. Commun. 181:286-292(1991).
- RA [14]  
RC PARTIAL SEQUENCE (SAPOSIN B), AND STRUCTURE OF CARBOHYDRATES.  
RC TISSUE=urine;  
RX MEDLINE=20032116; PubMed=10562467;  
RA Fluharty A.L., Lombard C., Louis A., Stevens R.L., Whitelegge J.P.,  
RA Waring A.J., To T., Fluharty C.B., Fluharty A.L.;  
RT "Preparation of the cerebroside sulfate activator (CSAct or saposin B)  
RT from human urine.";  
RL Mol. Genet. Metab. 68:391-403(1999).
- RA [15]  
RC STRUCTURE OF CARBOHYDRATE ON ASN-215.  
RX MEDLINE=21110404; PubMed=11180632;  
RA Pauli K.F., Johnson J., Kim M.J., To T., Whitelegge J.P.,  
RA Stevens R.L., Fluharty C.B., Fluharty A.L.;  
RT "Structure of the asparagine-linked sugar chains of porcine kidney and  
RT human urine cerebroside sulfate activator protein.";  
RL J. Mass Spectrom. 35:1416-1424(2000).
- RA [16]  
RC SAPOSIN D DISULFIDE BONDS.  
RX MEDLINE=99337688; PubMed=10406958;  
RA Tatti M., Salvioli R., Claffoni F., Pucci P., Andolfo A.,  
RA Amoresano A., Vaccaro A.M.;  
RT "Structural and membrane-binding properties of saposin D.";  
RL Eur. J. Biochem. 263:486-494(1999).
- RA [17]  
RC SAPOSIN B DISULFIDE BONDS.  
RX MEDLINE=2398398; PubMed=12510003;  
RA Ahn V.E., Pauli K.F., Whitelegge J.P., Higginson J., Fluharty A.L.,  
RA Prive G.G.;  
RT "Expression, purification, crystallization, and preliminary X-ray  
RT analysis of recombinant human saposin B.";  
RL Protein Expr. Purif. 27:186-193(2003).
- RA [18]  
RC MASS SPECTROMETRY.  
RC TISSUE=Urine;  
RX MEDLINE=99441404; PubMed=10510427;  
RA Pauli K.F., Whitelegge J.P., Higginson J., To T., Johnson J.,  
RA Krutchinsky A.N., Standing K.G., Waring A.J., Stevens R.L.,  
RA Fluharty C.B., Fluharty A.L.;  
RT "Cerebroside sulfate activator protein (Saposin B): chromatographic  
RT and electrospray mass spectrometric properties.";  
RL J. Mass Spectrom. 34:1040-1054(1999).
- RA [19]  
RC X-RAY CRYSTALLOGRAPHY (2.2 ANGSTROMS) OF 195-273, AND MUTAGENESIS OF  
RP ILE-240.  
RX MEDLINE=22406333; PubMed=12518053;  
RA Ahn V.E., Pauli K.F., Whitelegge J.P., Fluharty A.L., Prive G.G.;  
RT "Crystal structure of saposin B reveals a dimeric shell for lipid  
RT binding.";  
RL Proc. Natl. Acad. Sci. U.S.A. 100:38-43(2003).
- RA [20]  
RC REVIEW ON MLD VARIANTS.  
RX MEDLINE=95170731; PubMed=7866401;  
RA Gieselmann V., Zlotogora J., Harris A., Wenger D.A., Morris C.P.;  
RT "Molecular genetics of metachromatic leukodystrophy.";  
RL Hum. Mutat. 4:233-242(1994).
- RA [21]  
RC VARIANT MLD ILE-217.  
RX MEDLINE=90147748; PubMed=2302219;  
RA Rafi M.A., Zhang X.-L., Degala G., Wenger D.A.;  
RT "Detection of a point mutation in sphingolipid activator protein-1  
RT mRNA in patients with a variant form of metachromatic  
RT leukodystrophy.";  
RL Biochem. Biophys. Res. Commun. 166:1017-1023(1990).
- RA [22]  
RC SEQUENCE FROM N.A., AND VARIANT MLD ILE-217.  
RX MEDLINE=90207231; PubMed=2320574;

Query Match	Similarity	100.0%	Score 2759	DB 1	Length 524
Best Local	Similarity	100.0%	Pred. No. 1,86-173		
Matches	518	Conservative	0	Mismatches	0
				Indels	Gaps
QY	1	LASLILGALLAGVLLKECTRGSAAWCQWVKVASDCGAVKHCQLQTVWNNKPYKSLPCDIC	60		
DB	7	LASLILGALLAGVLLKECTRGSAAWCQWVKVASDCGAVKHCQLQTVWNNKPYKSLPCDIC	66		
QY	61	KDVVTAAGDMLKDNATEEELVYLEKTCMDLPRKPNMSASCKEIVDSYLPVLIIDIIKGENS	120		
DB	67	KDVVTAAGDMLKDNATEEELVYLEKTCMDLPRKPNMSASCKEIVDSYLPVLIIDIIKGENS	126		
QY	121	RPGVECSALNLCESQKHLAEINHKOLESNKIPELDMEVVAAPPMANIPILLIYQDGR	180		
DB	127	RPGVECSALNLCESQKHLAEINHKOLESNKIPELDMEVVAAPPMANIPILLIYQDGR	186		
QY	181	SKPGKNDGVQDCICQWVTDIQTAVRTSTVQALVEHVEKCECRLGSGMADICKNYTS	240		
DB	187	SKPGKNDGVQDCICQWVTDIQTAVRTSTVQALVEHVEKCECRLGSGMADICKNYTS	246		
QY	241	QYSEIAIQMMHMQKEICALVGFCEVKEPMQTLVPKVASKNVIPALAEVPIKXHE	300		
DB	247	QYSEIAIQMMHMQKEICALVGFCEVKEPMQTLVPKVASKNVIPALAEVPIKXHE	306		
QY	301	VPAKSDVYCEVCEPFLVKEVTKLIDNNKTEKELIDPAFDKRCSTLPSLSCEQEVVDYGS	360		
DB	307	VPAKSDVYCEVCEPFLVKEVTKLIDNNKTEKELIDPAFDKRCSTLPSLSCEQEVVDYGS	366		
QY	361	SILSILTEEVSPELVCSMLHLCSGTRLPALTVHTVPKQGGFCVCKLWGLDRLNLEKN	420		
DB	367	SILSILTEEVSPELVCSMLHLCSGTRLPALTVHTVPKQGGFCVCKLWGLDRLNLEKN	426		
QY	421	STKQEIILALEKGSFLLPDYQKCDQVAAEYEPVILIEILVEVMDPSFVCLKIGACPSAH	480		
DB	427	STKQEIILALEKGSFLLPDYQKCDQVAAEYEPVILIEILVEVMDPSFVCLKIGACPSAH	486		
QY	481	KPLIGTEKCIWGPSYMCNTEFTAAOCNVAEHCKRHWN	518		
DB	487	KPLIGTEKCIWGPSYMCNTEFTAAOCNVAEHCKRHWN	524		

RESULT 2

SAP BOVIN

ID	SAP BOVIN	STANDARD	PRT	525 AA
AC	P26779	Q9N2G4		
DT	01-AUG-1992	(Rel. 23, Created)		
DT	16-OCT-2001	(Rel. 40, Last sequence update)		
DT	16-OCT-2001	(Rel. 40, Last annotation update)		
DE	Proactivator polypeptide precursor [Contains: Saposin A (Protein A); Saposin B (Sphingolipid activator protein 1) (SAP-1) (Cerebroside sulfate activator) (CSact) (Dispersin) (Sulfatide) (Gm1 activator); Saposin C (Co-beta-glucosidase) (A1 activator) (Glucosylceramidase activator) (Sphingolipid activator protein 2) (SAP-2); Saposin D (Protein C) (Component C)].			
DE	PSAP.			
OS	Bos taurus (Bovine).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae; Bovidae; Bovinae; Bos.			
OC	NCBI_TaxID=9913;			
OX	NCBI_TaxID=9913;			
RN	SEQUENCE FROM N.A., AND VARIANTS.			
RC	TISSUE=Mammary gland;			
RA	Azuma N., Yoshida K.;			
RT	"RT-PCR cloning of bovine prosaposin.";			
RL	Submitted (JAN-2000) to the EMBL/GenBank/DBJ databases.			
RP	SEQUENCE OF 312-391.			
RC	TISSUE=Spleen;			
RX	MEDLINE=92207994; PubMed=1554743;			
RA	Sano A., Mizuno T., Kondoh K., Himeeno T., Ueno S.-I., Kakimoto Y., Morita N.;			

	"asposin-C from bovine spleen; complete amino acid sequence and relation between the structure and its biological activity.";
Rt	Biochim. Biophys. Acta 1120:75-80(1992).
-I-	FUNCTION: The lysosomal degradation of sphingolipids takes place by the sequential action of specific hydrolases. Some of these enzymes require specific low-molecular mass, non-enzymic proteins: the sphingolipids activator proteins (coproteins) (By similarity).
-I-	FUNCTON: Saposin A and seposin C stimulate the hydrolysis of galactosylceramide by beta-galactosylceramidase (EC 3.2.1.45) and saposin B apparently acts by combining with the enzyme and acidic lipid to form an activated complex, rather than by solubilizing the substrate.
-I-	FUNCTON: Saposin B stimulates the hydrolysis of galacto-cerebroside sulfate by arylsulfatase A (EC 3.1.6.8), GM1 gangliosidosis by beta-galactosidase (EC 3.2.1.23) and galabiosylceramide by alpha-galactosidase A (EC 3.2.1.22). Saposin B forms a solubilizing complex with the substrates of the sphingolipid hydrolases (By similarity).
-I-	FUNCTON: Saposin D is a specific sphingomyelin phosphodiesterase activator (EC 3.1.4.12) (By similarity).
-I-	SUBUNIT: Saposin B is a homodimer (By similarity).
-I-	SUBCELLULAR LOCATION: Lysosomal.
-I-	PRT: This precursor is proteolytically processed to 4 small peptides, which are similar to each other and are sphingolipid hydrolase activator proteins (By similarity).
-I-	SIMILARITY: Contains 2 saposin A-type domains.
-I-	SIMILARITY: Contains 4 saposin B-type domains.
-- -- -- -- --	
Cc	This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outpostation - the European Bioinformatics institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).
Cc	-- -- -- -- --
Df	EMBL; AB036791; BAA95677.1; ..
Dr	InterPro; IPR003119; SAPA.
DR	InterPro; IPR007856; SAPB_1.
Dr	InterPro; IPR008138; SAPB_2.
DR	InterPro; IPR008140; SapB_sub.
DR	InterPro; IPR008373; Saposin.
DR	InterPro; IPR008139; SaposinB.
DR	pfam; PF02119; SAPA; 2.
DR	pfaam; PF05184; SAPB_1; 4.
DR	pfaam; PF03489; SAPB_2; 4.
DR	PRINTS; PR01797; SAPOSIN.
DR	Prodrom; PD001732; SapB_sub; 3.
DR	SMART; SMART0162; SAPA; 2.
DR	SMART; SMART0118; SAPB; 4.
KW	Signal, Glycoprotein, Lysosome, Sphingolipid metabolism; Repeat.
FT	SIGNAL 1       16
FT	PROPEP 17      58
FT	CHAIN 60     142
FT	PROPEP 144     195
FT	CHAIN 196     275
FT	PROPEP 277     310
FT	CHAIN 312     392
FT	PROPEP 393     404
FT	CHAIN 406     487
FT	PROPEP 489     525
FT	DOMAIN 21      54
FT	DOMAIN 59     142
FT	DOMAIN 194     276
FT	DOMAIN 312     393
FT	DOMAIN 406     487
FT	DOMAIN 492     525
FT	DISULFID 63     138
FT	DISULFID 66     132
FT	DISULFID 94     106
FT	DISULFID 199     272
FT	DISULFID 202     266
	SAPOSIN-LIKE TYPE A 1. SAPOSIN-LIKE TYPE B 1. SAPOSIN-LIKE TYPE B 2. SAPOSIN-LIKE TYPE B 3. SAPOSIN-LIKE TYPE B 4. SAPOSIN-LIKE TYPE A 2. BY SIMILARITY. BY SIMILARITY. BY SIMILARITY. BY SIMILARITY. BY SIMILARITY.

FT DISULFID 231 242 BY SIMILARITY.  
 FT DISULFID 316 389 BY SIMILARITY.  
 FT DISULFID 319 383 BY SIMILARITY.  
 FT DISULFID 347 358 BY SIMILARITY.  
 FT DISULFID 410 483 BY SIMILARITY.  
 FT DISULFID 413 477 BY SIMILARITY.  
 FT DISULFID 441 452 BY SIMILARITY.  
 FT CARBOHYD 80 80 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 101 101 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 216 216 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 333 333 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 427 427 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT VARIANT 127 127 H -> R.  
 FT VARIANT 260 263 MOCK -> IRIR.  
 FT VARIANT 317 317 E -> Q (IN REF. 2).  
 FT CONFLICT 367 367 R -> S (IN REF. 1).  
 SQ SEQUENCE 525 AA; 58120 MW; 293AF0FB9C4F9A99 CRC64;

Query Match 87.2%; Score 2405.5; DB 1; Length 525;  
 Best Local Similarity 84.6%; Pred. No. 2.5e-150;  
 Matches 439; Conservative 43; Mismatches 36; Indels 1; Gaps 1;

QY 1 LASLIGALAGAVLGLKCTGRSAWCVQVKTASGAVGCLQTVNKPVKSPIC 60  
 DB 7 LASLIGALASPVGLGRLCTGRSAWCVQVKTADGAVOHLQTVNKPVKSPIC 66  
 QY 61 KDVTAAGDMLKDNATEERILVYLEKTCVPLPKPMASCKEIVSYLPVLDIKGMS 120  
 DB 67 KDVITRAGGLKDNATEERILVYLEKTCVPLPKPMASCKEIVSYLPVLDIKGMS 126  
 QY 121 RRGVCSALNLCESIQKHLAEINHKQLESNKIPELDMTEVAPPMANIPLLYPQDGP 180  
 DB 127 HGEVCSALNLCESIQKHLAEINHKQLESNQPILDMAEVAPPMANIPFLYPQDSH 186  
 QY 181 SKPQKRD-NGDPQDCQICQVAVTDTQVAVTNSFVQVAVENHKEEDRIGPMDICNYI 239  
 DB 187 SKPQKRGANVCQCCICQVAVTDTQVAVTNSFVQVAVENHKEEDRIGPMDICNYI 246  
 QY 240 SQYSIAIOMMMHMQPKICALVGFCDVCEKEMQTVLPAYAKVKNVPALEVEPIKCH 299  
 DB 247 NQYSEVALIOMVMMHMQPKICVLAGFCDEKEMPKTVLPAYVSENVIPALGLVEPIKD 306  
 QY 300 EVPAKSDVYCEVCEFLVEVYTKLIDNKTKEIIDLDFKWCSPKLSISECOEVDYTG 359  
 DB 307 PAPAAADYICVCEVVEVAVAKLIDNKTKEIIDLDFKWCSPKLSISECOEVDYTG 366  
 QY 360 SSILSILEEVSPELVCSMLHLCSGTRLPALTVAHYTPQKQGFCEVCKLVGYDLNLEK 419  
 DB 367 RSILSILDEASPELVCSMLHLCSSRGRLPATVAVMPKQDGFCEVCKLVGYDLNLEK 426  
 QY 420 NSTQOEIILAEKGCSPDPDYQKQDQFVAEYEPVLEIILVEVMDSPFVCLKIGACPSA 479  
 DB 427 NSTQOEIILAEKGCSPDPDYQKQDQFVAEYEPVLEIILVEVMDSPFVCLKIGACPSA 486  
 QY 480 HKPILGTEKCIWSPSYMCONMETAAQCAVNHCHGRHWAN 518  
 DB 487 HKPILGAKCVMGSPSYMCONMESALCAVNHCHGRHWAN 525

RESULT 3  
 SAP\_RAT STANDARD; PRT; 554 AA.  
 AC P10560; 062841; 064190;  
 DT 01-JUL-1989 (Rel. 11, Created)  
 DT 01-JUL-1989 (Rel. 11, Last sequence update)  
 DE 15-MAR-2004 (Rel. 43, Last annotation update)  
 OS sulfated glycoprotein 1 precursor (SGP-1) (Prosaposin).  
 GN PSAP OR SGPI.  
 OS Rattus norvegicus (Rat).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.  
 OX NCBI\_Taxid=10116;  
 RN [1]

RP SEQUENCE FROM N.A. AND PARTIAL SEQUENCE.  
 RC TISSUE=Sertoli cells;  
 RA MEDLINE=89000647; PubMed=3048385;  
 RX Collard M.W., Sylvester S.R., Tsuruta J.K., Griswold M.D.,  
 RT "Biosynthesis and molecular cloning of sulfated glycoprotein 1  
 RT 70-kilodalton precursor to sulfatide/GM1 activator",  
 RL Biochemistry 27:4557-4564(1988).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Testis;  
 RX MEDLINE=96128541; PubMed=8573994;  
 RA Morales C.R., El-Alfy M., Zhao Q., Igodura S.A.,  
 RT "Molecular role of sulfated glycoprotein-1 (SGP-1/prosaposin) in  
 RT Sertoli cells",  
 RL Histol. Histopathol. 10:1023-1034(1995).  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Testicle;  
 RX MEDLINE=96175245; PubMed=8601692;  
 RA Morales C.R., El-Alfy M., Zhao Q., Igodura S.A.,  
 RT "Expression and tissue distribution of rat sulfated glycoprotein-1  
 RT (prosaposin).",  
 RL J. Histochem. Cytochem. 44:327-337(1996).  
 CC -1- SUBCELLULAR LOCATION: Extracellular.  
 CC -1- SIMILARITY: Contains 2 saposin A-type domains.  
 CC -1- SIMILARITY: Contains 4 saposin B-type domains.  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL; M19936; AAA42136.1; -  
 DR EMBL; S81353; AAB36042.2; -  
 DR EMBL; S81373; AAB36235.2; -  
 DR PIR; A28716; A28716.  
 DR InterPro; IPR003119; Sapa.  
 DR InterPro; IPR007856; Sapa\_1.  
 DR InterPro; IPR008138; Sapa\_2.  
 DR InterPro; IPR008140; Sapa\_sub.  
 DR InterPro; IPR008373; Saposin.  
 DR InterPro; IPR008139; Saposinb.  
 DR Pfam; PF02199; SAPA; 2.  
 DR Pfam; PF05184; Sapa\_1; 3.  
 DR Pfam; PF03489; Sapa\_2; 4.  
 DR PRINTS; PR01797; SAPOSIN.  
 DR ProDom; PD001732; Sapa\_sub; 3.  
 DR SMART; SM00162; SAPA; 2.  
 DR SMART; SM00118; SApB; 4.  
 KW Signal; Glycoprotein; Repeat.  
 FT SIGNAL 1 16  
 FT CHAIN 17 554 SULFATED GLYCOPROTEIN 1.  
 FT DOMAIN 21 54 SAPOSIN-LIKE TYPE A 1.  
 FT DOMAIN 61 138 SAPOSIN-LIKE TYPE B 1.  
 FT DOMAIN 193 274 SAPOSIN-LIKE TYPE B 2.  
 FT DOMAIN 310 391 SAPOSIN-LIKE TYPE B 3.  
 FT DOMAIN 435 516 SAPOSIN-LIKE TYPE B 4.  
 FT DOMAIN 521 554 SAPOSIN-LIKE TYPE A 2.  
 FT DISULFID 63 138 BY SIMILARITY.  
 FT DISULFID 66 132 BY SIMILARITY.  
 FT DISULFID 94 106 BY SIMILARITY.  
 FT DISULFID 197 270 BY SIMILARITY.  
 FT DISULFID 200 264 BY SIMILARITY.  
 FT DISULFID 229 240 BY SIMILARITY.  
 FT DISULFID 314 387 BY SIMILARITY.  
 FT DISULFID 317 381 BY SIMILARITY.  
 FT DISULFID 345 356 BY SIMILARITY.  
 FT DISULFID 439 512 BY SIMILARITY.  
 FT DISULFID 442 506 BY SIMILARITY.

```

FT DISULFID 470 481 BY SIMILARITY.
FT CARBOHYD 80 80 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 214 214 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 331 331 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 456 456 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 115 115 P -> L (IN REF. 2 AND 3).
FT CONFLICT 299 299 D -> E (IN REF. 2).
FT CONFLICT 462 462 I -> V (IN REF. 3).
FT CONFLICT 527 527 W -> R (IN REF. 3).
FT CONFLICT 536 536 S -> M (IN REF. 3).
SQ SEQUENCE 554 AA; 61123 MW; DFE3FA3A0520C6B CRC64;

Query Match 71.5%; Score 1974; DB 1; Length 554;
Best Local Similarity 66.1%; Pred. No. 4,8e-122;
Matches 363; Conservative 77; Mismatches 77; Indels 32; Gaps 3;

QY 1 LASLIGAAAGPVGLKECTRGSAVQNVKTASDCGAVKCLQTVNKPVKSLPCDIC 60
DB 7 LASLIVLTALTSFVQPKICSGGSAVVCRRDKTAVDCRAVKHCQQVWMSKPTAKSLPCDIC 66
QY 61 KDVTAAAGDMKDNATEEELIVLEKTCQDMKPKPMMSACEIYDSYLPVLLDIKIGMS 120
DB 67 KTVVTEAENLTKDNATEEELIVLEKTCQDMKPKPMMSACEIYDSYLPVLLDIKIGMS 126
QY 121 RPEVCALNLCESLQKLAELNHOKESENKIPELDMTEVVAPEMANIPILLYPODGR 180
DB 127 NPEVCALNLCESLQKLAELNHOKESENKIPELDMTEVVAPEMANIPILLYPODGR 185
QY 181 SKPOKNDGVQCQDCIQVNTIQTAVRTNSTFVQALVEHVKEECRLGPGMADICKNYIS 240
DB 186 SOPKKNEDVQCQDCMKLVTDIQTAVRTNSTFVQALVEHVKEECRLGPGVADICKNYVD 245
QY 241 QYSEIATLQMMHMOKEICALVGFCEVKEPMQTLVPAKASKAVIPALVEIKKHE 300
DB 246 QYSEVAVQMMHMOKEICVAVGFCEVKEPMQTLVPAKASKAVIPALVEIKKHE 305
QY 301 VPAKSDVCEVCEFLVKEVTKLIDNNKTEKELIDAFDKMCSLPKSLSECOEVDYTGSS 360
DB 306 IQQNVITICQVQALVWRKLSLILINNATEELIKLSKACSLILPAPASTKCOEVLVTEGP 365
QY 361 SILSILLESVPELVCSMLHLCSG-----TRLP-----A 389
DB 366 SILDVLMEVNPVNFGLGVISLCSANPNLVGLEQPAALIVSALPREPAPKQPEBPQKSA 425
QY 426 LRAHVPQKNGGFCVCKKLVYILHNLKXSTKEIILAALEKGSFLPDPYQKOCDEFV 485
DB 450 AVEPVLLLEIVVNDPSFVCLKIGACPSAHPFLGTGECIMGSPSWCONTEATAQCNV 509
QY 486 AVEPVLLLEIVVNDPSFVCLKIGACPSAHPFLGTGECIMGSPSWCONTEATAQCNV 545
QY 510 EHCKRHVNN 518
DB 546 DHCKRHVNN 554

RESULT 4
SAP_MOUSE STANDARD; PRT; 557 AA.
ID 061307; 060861; 064219;
AC 15-DEC-1998 (Rel. 37, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DE Sulfated glycoprotein 1 precursor (SGP-1) (Prosaposin).
GN PSAP OR SGP1.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
CX NCBI_Taxid=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Liver;

```

```

RX MEDLINE=92272718; PubMed=1590788;
RA Tsuda M., Sakiyama T., Endo H., Kitagawa T.;
RT "The primary structure of mouse saposin.";
RL Biochem. Biophys. Res. Commun. 184:1266-1272(1992).
RN (2)
RP SEQUENCE FROM N.A.
RX MEDLINE=94272117; PubMed=8003952;
RA Sprecher-Levy H., Orr-Urtreger A., Lonal P., Horowitz M.;
RT "Murine prosaposin: expression in the reproductive system of a gene
RT implicated in human genetic disease.";
RL Cell. Mol. Biol. 40:233-233(1994).
RN (3)
RP SEQUENCE FROM N.A.
RX MEDLINE=96084310; PubMed=8565332;
RA Cao Q.P., Crain W.R.;
RT "Expression of SGP-1 mRNA in preimplantation mouse embryos.";
RL Dev. Genet. 17:263-271(1995).
RN (4)
RP SEQUENCE FROM N.A.
RC STRAIN=BALB/c; TISSUE=Liver;
RA Zhao Q.Q., Hay N.N., Morales C.R.;
RL Submitted (May-1996) to the EMBL/GenBank/DBJ databases.
CC -1- SUBCELLULAR LOCATION: Extracellular.
CC -1- SIMILARITY: Contains 2 saposin A-type domains.
CC -1- SIMILARITY: Contains 4 saposin B-type domains.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL, S36200; AAB22175.1; -
DR EMBL, S71616; AAB31059.1; -
DR EMBL, U27340; AAB92567.1; -
DR EMBL, U57999; AAB02695.1; -
DR PIR, UH0604; UH0604.
DR MGI, MGI:97783; Psap.
DR InterPro; IPR003119; SAPA.
DR InterPro; IPR007856; SAPB 1.
DR InterPro; IPR008138; SAPB 2.
DR InterPro; IPR008139; SAPB 2.
DR InterPro; IPR008139; SAPB 2.
DR Pfam; PF02199; SAPA; 2.
DR Pfam; PF05184; SAPB 1; 4.
DR Pfam; PF03489; SAPB 2; 4.
DR PRINTS; PR01797; SAPOSIN.
DR ProDom; PD001732; SapB_sub; 3.
DR SMART; SM00162; SAPA; 2.
DR SMART; SM00118; SAPB; 4.
KW Signal; Glycoprotein; Repeat.
FT SIGNAL 1 16
FT CHAIN 17 557 BY SIMILARITY.
FT DOMAIN 21 54 SAPOSIN-LIKE TYPE A 1.
FT DOMAIN 59 142 SAPOSIN-LIKE TYPE B 1.
FT DOMAIN 193 277 SAPOSIN-LIKE TYPE B 2.
FT DOMAIN 313 394 SAPOSIN-LIKE TYPE B 3.
FT DOMAIN 438 519 SAPOSIN-LIKE TYPE B 4.
FT DOMAIN 524 557 SAPOSIN-LIKE TYPE A 2.
FT DISULFID 63 138 BY SIMILARITY.
FT DISULFID 66 132 BY SIMILARITY.
FT DISULFID 94 106 BY SIMILARITY.
FT DISULFID 197 273 BY SIMILARITY.
FT DISULFID 200 267 BY SIMILARITY.
FT DISULFID 229 240 BY SIMILARITY.
FT DISULFID 317 390 BY SIMILARITY.
FT DISULFID 320 384 BY SIMILARITY.
FT DISULFID 348 359 BY SIMILARITY.
FT DISULFID 442 515 BY SIMILARITY.
FT DISULFID 445 509 BY SIMILARITY.

```

FT DISULFID 473 484 BY SIMILARITY.  
 FT CARBOHYD 80 80 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 214 214 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 334 334 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 459 459 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CONFLICT 83 83 Q -> E (IN REF. 2).  
 FT CONFLICT 158 158 I -> V (IN REF. 3).  
 FT CONFLICT 160 160 MISSING (IN REF. 2).  
 FT CONFLICT 171 172 MS -> SA (IN REF. 3).  
 FT CONFLICT 244 244 V -> L (IN REF. 2).  
 FT CONFLICT 254 254 M -> I (IN REF. 3).  
 FT CONFLICT 255 255 L -> W (IN REF. 3).  
 FT CONFLICT 260 260 MISSING (IN REF. 3).  
 FT CONFLICT 307 307 N -> D (IN REF. 2).  
 FT CONFLICT 322 322 F -> L (IN REF. 2).  
 FT CONFLICT 349 350 AL -> GV (IN REF. 1).  
 FT CONFLICT 367 367 G -> D (IN REF. 3).  
 FT CONFLICT 370 370 L -> Q (IN REF. 2).  
 FT CONFLICT 372 373 I -> D (IN REF. 3).  
 FT CONFLICT 391 391 A -> T (IN REF. 3).  
 FT CONFLICT 393 393 R -> L (IN REF. 3).  
 FT CONFLICT 406 406 A -> R (IN REF. 2 AND 3).  
 FT CONFLICT 430 430 P -> R (IN REF. 2).  
 FT CONFLICT 445 445 C -> F (IN REF. 3).  
 FT CONFLICT 448 448 L -> P (IN REF. 4).  
 SQ SEQUENCE 557 AA; 61422 MM; 134593E049B35E CRC64;

Query Match 69.5%; Score 1916.5; DB 1; Length 557;  
 Best Local Similarity 63.5%; Pred. No. 2.8e-116;  
 Matches 350; Conservative 78; Mismatches 88; Indels 35; Gaps 3;

QY 2 ASLIGALAGPVLGKECTRGASVAVCCGAVYHGLQTVWANKPYKSPDCDCK 61  
 DB 8 ASLIGALAGPVLGKECTRGASVAVCCGAVYHGLQTVWANKPYKSPDCDCK 67  
 QY 62 DVATAGDMLKDNATEEELVLEKTCQDMLPKPMNASCKEIVSYLPIVLIIDIKGEMSR 121  
 DB 68 TVTETAGMLKDNATEEELVLEKTCQDMLPKPMNASCKEIVSYLPIVLIIDIKGEMSN 127  
 QY 122 PGEVVSANLCEGLOKHLAEINHOLESNKI PELDMTEVVAFPMANIPLLLYFODGPRS 181  
 DB 128 PGEVVSANLCEGLOKHLAEINHOLESNKI PELDMTEVVAFPMANIPLLLYFODGPRS 186  
 QY 182 KPOPKDNGVODCICOWMTDIOTARTSTVQALVEHVEKECRLGPMADICKYISQ 241  
 DB 187 QPQPMANEDVODCICOWMTDIOTARTSTVQALVEHVEKECRLGPMADICKYISQ 246  
 QY 242 YSEIAIQMMEM--QPKICALVGFCEVEKEMQTLVPAKVASKVYIPALTEVEPIKK 298  
 DB 247 YSEVCGVQMLMHOQOQPKICVLAGFCNEVAVKPMKTLVPAETITIKUILPALMEMDYEQ 306  
 QY 299 HEVPAKSVVCEVEBELVKEVTKLIDNNKTEKELIDAFDMKCSKLPKSLSEECQEVVDY 358  
 DB 307 NLVQNHANVITLCQTCQFVWNRKSELIVNNAIEELLVKLSNACALLPDPARTKCEVVGTF 366  
 QY 359 GSSILSLILEEVSPELVCSMLHLCG-----TRL 387  
 DB 367 GSPSLDFIHIEVNVSSLCGVIGLCARPVELVALEQAPALVSKLKEPPLPPQAPQPK 426  
 QY 388 PALTVHTVTPKDGFCVCEVKLVYLDRLNLEKSTKQEIILALEKSGSPFLPDYQKQCD 447  
 DB 427 SALPAHVPPQKNGFCVCEVKLVYLEHNLKSTKEELILALEKSGSPFLPDYQKQCD 466  
 QY 448 FVAEYEPVLEILIEVWMDPSVFLKIGACPSAHNPILGTEKCTWBSYVQNTETAQCN 507  
 DB 487 FVAEYEPVLEILIEVWMDPSVFLKIGACPSAHNPILGTEKCTWBSYVQNTETAQCN 546  
 QY 508 AVEHCCKHWN 518  
 DB 547 AVDHCKHWN 557

RESULT 5

SAP\_CHICK  
 ID SAP\_CHICK STANDARD; PRT; 518 AA.  
 AC 013035;  
 DT 16-OCT-2001 (Rel. 40, Created)  
 DT 16-OCT-2001 (Rel. 40, Last sequence update)  
 DT 28-FEB-2003 (Rel. 41, Last annotation update)  
 DE Proactivator polypeptide precursor [Contains: Saposin A; Saposin B; Saposin C; Saposin D].  
 GN PSAP.  
 OS Gallus gallus (Chicken).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;  
 OC Gallus.  
 OX NCBI\_TaxID=9031;  
 RN [1]  
 RP SEQUENCE FROM N.A., AND SEQUENCE OF 194-203.  
 RC TISSUE=Brain, and Liver;  
 RX MEDLINE=98129745; PubMed=9461526;  
 RA Azuma N., Seo H.-C., Iie O., Pu Q., Gould R.M., Hiraiwa M., Burt D.W.,  
 RA Paton I.R., Morris D.R., O'Brien J.S., Kishimoto Y.,  
 RT "Cloning, expression and map assignment of chicken prosaposin.";  
 RL Biochem. J. 330:321-327(1998).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RA Altman N., Horowitz M.;  
 RL Submitted (NOV-1998) to the EMBL/GenBank/DBJ databases.  
 CC -1- FUNCTION: The lysosomal degradation of sphingolipids takes place by the sequential action of specific hydrolases. Some of these enzymes require specific low-molecular mass, non-enzymic proteins: the sphingolipids activator proteins (coproteins) (By similarity).  
 CC -1- FUNCTION: Saposin A and saposin C stimulate the hydrolysis of glucosylceramide by beta-glucosylceramidase (EC 3.2.1.45) and galactosylceramide by beta-galactosylceramidase (EC 3.2.1.46).  
 CC Saposin C apparently acts by combining with the enzyme and acidic lipid to form an activated complex, rather than by solubilizing the substrate (By similarity).  
 CC -1- FUNCTION: Saposin B stimulates the hydrolysis of galactogangliosides by beta-galactosidase (EC 3.2.1.23) and globotriaosylceramide by alpha-galactosidase A (EC 3.1.6.8), GM1 saposin B forms a solubilizing complex with the substrates of the sphingolipid hydrolases (By similarity).  
 CC -1- FUNCTION: Saposin D is a specific sphingomyelin phosphodiesterase activator (EC 3.1.4.12) (By similarity).  
 CC -1- SUBUNIT: Saposin B is a homodimer (By similarity).  
 CC -1- SUBCELLULAR LOCATION: Lysosomal (By similarity).  
 CC -1- PTM: This precursor is proteolytically processed to 4 small peptides, which are similar to each other and are sphingolipid hydrolase activator proteins (By similarity).  
 CC -1- SIMILARITY: Contains 2 saposin A-type domains.  
 CC -1- SIMILARITY: Contains 4 saposin B-type domains.  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).  
 CC  
 DR EMBL AB004471; BAA19914.1; -;  
 DR EMBL AF108656; AAF05899.1; -;  
 DR InterPro: IPR003119; SABA.  
 DR InterPro: IPR007856; SABA\_1.  
 DR InterPro: IPR008138; SABA\_2.  
 DR InterPro: IPR008140; SABA\_sub.  
 DR InterPro: IPR008373; Saposin.  
 DR InterPro: IPR008139; SaposinB.  
 DR Pfam PF02199; SABA; 2.  
 DR Pfam PF05184; SABA\_1; 4.  
 DR Pfam PF03489; SABA\_2; 4.  
 DR PRINTS PR01797; Saposin.  
 DR ProDom PD001732; SABA\_sub; 3.

KW Signal; Glycoprotein; Lysosome; Sphingolipid metabolism; Repeat;  
 KW GM2-gangliosidosis.  
 FT SIGNAL 1 17 POTENTIAL.  
 FT PROPER 18 60  
 FT CHAIN 61 143 SAPOSIN A.  
 FT PROPER 145 193  
 FT CHAIN 194 276 SAPOSIN B.  
 FT PROPER 278 305  
 FT CHAIN 307 387 SAPOSIN C.  
 FT PROPER 389 398  
 FT CHAIN 399 480 SAPOSIN D.  
 FT PROPER 482 518  
 FT DOMAIN 22 55 SAPOSIN-LIKE TYPE A 1.  
 FT DOMAIN 60 143 SAPOSIN-LIKE TYPE B 1.  
 FT DOMAIN 193 277 SAPOSIN-LIKE TYPE B 2.  
 FT DOMAIN 307 388 SAPOSIN-LIKE TYPE B 3.  
 FT DOMAIN 399 480 SAPOSIN-LIKE TYPE B 4.  
 FT DOMAIN 485 518 SAPOSIN-LIKE TYPE A 2.  
 FT DISULFID 64 139 BY SIMILARITY.  
 FT DISULFID 67 133 BY SIMILARITY.  
 FT DISULFID 95 107 BY SIMILARITY.  
 FT DISULFID 197 273 BY SIMILARITY.  
 FT DISULFID 200 267 BY SIMILARITY.  
 FT DISULFID 229 240 BY SIMILARITY.  
 FT DISULFID 311 384 BY SIMILARITY.  
 FT DISULFID 314 378 BY SIMILARITY.  
 FT DISULFID 342 353 BY SIMILARITY.  
 FT DISULFID 403 476 BY SIMILARITY.  
 FT DISULFID 406 470 BY SIMILARITY.  
 FT DISULFID 434 445 BY SIMILARITY.  
 FT CARBOHYD 81 81 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 214 214 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 328 328 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 420 420 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CONFLICT 94, 94 R -> T (IN REF. 2).  
 FT CONFLICT 486 486 B -> D (IN REF. 2).  
 FT SEQUENCE 518 AA; 57601 MW; B803000E891C3963 CRC64;  
 Query Match 60.8%; Score 1676.5; DB 1; Length 518;  
 Beot Local Similarity 57.6%; Pred. No. 1.3e-10;  
 Matches 300; Conservative 96; Mismatches 112; Indels 13; Gaps 5;

QY 478 SAKRPLCTEKCIWGPSYWCNTERRAQCNAVEHCKRHVN 518  
 DB 478 AKKPLLEDGACVWGPYCWKMETPAQCNAVDHCRHVN 518  
 RESULT 6  
 SAP\_PIG STANDARD; PRT; 80 AA.  
 AC P81405;  
 DT 15-DEC-1998 (Rel. 37, Created)  
 DT 16-OCT-2001 (Rel. 40, Last sequence update)  
 DT 28-FEB-2003 (Rel. 41, Last annotation update)  
 DE Saposin B (Cerebroside sulfate activator) (CS-ACT) (Non-specific activator) (Sphingolipid activator protein 1) (SAP-1).  
 OS Sus scrofa (Pig).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Euteria; Cetartiodactyla; Suina; Suidae; Sus.  
 OX NCBI\_TaxId=9823;  
 RN [1]  
 RP SEQUENCE OF 1-79.  
 RC TISSUE=Kidney;  
 RX MEDLINE=93229506; PubMed=8471613;  
 RA Stevens R.L., Fauli K.F., Conklin K.A., Green B.N., Fluharty A.L.;  
 RT "Porcine cerebroside sulfate activator: further structural characterization and disulfide identification.";  
 RL Biochemistry 32:4051-4059(1993).  
 RN [2]  
 RP SEQUENCE OF 1-64.  
 RC TISSUE=Kidney;  
 RX MEDLINE=9222651; PubMed=1562358;  
 RA Fluharty A.L., Katona Z., Meek W.B., Frel K., Fowler A.V.;  
 RT "The cerebroside sulfate activator from pig kidney: purification and molecular structure.";  
 RL Biochem. Med. Metab. Biol. 47:66-85(1992).  
 RN [3]  
 RP STRUCTURE OF CARBOHYDRATE ON ASN-21.  
 RX MEDLINE=21110404; PubMed=11180632;  
 RA Fauli K.F., Johnson J., Kim M.J., To T., Whitelegge J.P., Stevens R.L., Fluharty C.B., Fluharty A.L.;  
 RT "Structure of the asparagine-linked sugar chains of porcine kidney and human urine cerebroside sulfate activator protein.";  
 RL J. Mass Spectrom. 35:1416-1424(2000).  
 RN [4]  
 RP MASS SPECTROMETRY.  
 RC TISSUE=Kidney;  
 RX MEDLINE=99441404; PubMed=10510427;  
 RA Fauli K.F., Whitelegge J.P., Higginson J., To T., Johnson J., Krutichinsky A.N., Standing K.G., Waring A.J., Stevens R.L., Fluharty C.B., Fluharty A.L.;  
 RT "Cerebroside sulfate activator protein (Saposin B): chromatographic and electrospray mass spectrometric properties.";  
 RL J. Mass Spectrom. 34:1040-1054(1999).  
 CC -!- FUNCTION: Saposin B stimulates the hydrolysis of galacto-cerebrosides by beta-galactosidase A (EC 3.2.1.23) and globotriaosylceramide by alpha-galactosidase A (EC 3.2.1.22). Saposin B forms a solubilizing complex with the substrates of the sphingolipid hydrolases.  
 CC Saposin B is a homodimer (by similarity).  
 CC -!- SUBUNIT: Saposin B is a homodimer (by similarity).  
 CC -!- PFM: The one residue extended Saposin B-Val is only found in a minority of the chains.  
 CC -!- SIMILARITY: Contains 1 saposin B-type domain.  
 DR GlycositeDB; P81405; -.  
 DR InterPro; IPR007856; SAPB\_1.  
 DR InterPro; IPR008138; SAPB\_2.  
 DR InterPro; IPR008373; Saposin.  
 DR InterPro; IPR008139; SaposinB.  
 DR Pfam; PF05184; SAPB\_1; 1.  
 DR Pfam; PF03489; SAPB\_2; 1.  
 DR PRINTS; PR01797; SAPOSIN.  
 DR SMART; SM00118; SAPB; 1.  
 KW Glycoprotein; Sphingolipid metabolism.



```

FT CHAIN 1 79 SAPOSIN B:
FT CHAIN 1 80 SAPOSIN B-VAL.
FT DOMAIN 1 80 SAPOSIN-LIKE TYPE B.
FT DISULFID 4 77
FT DISULFID 7 71
FT DISULFID 36 47
FT CARBOHYD 21 21
SQ SEQUENCE 80 AA; 8949 MW; EF7BA29B63E789C CRC64;
N-LINKED (GLCNAC. . .) (COMPLEX).
/FTID-CAR 000177.

Query Match 14.2%; Score 393; DB 1; Length 80;
Best Local Similarity 88.8%; Pred. No. 1.4e-19;
Matches 71; Conservative 5; Mismatches 4; Indels 0; Gaps 0;

Dy 189 GDVCDICQWVTTIQAVRTNSTFVQALVEHKECDRLGPGMADCKRYISGYSIAIQ 248
1 GDVCDICQWVTTIQAVRTNSTFVQALVEHKECDRLGPGMADCKRYISGYSIAIQ 60
Dy 249 MAMHMQPKETICALVGFCEV 268
61 MAMHMQPKETICALVGFCEV 80

RESULT 7
PSPB_MOUSE STANDARD; PRT; 377 AA.
AC P50405;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DE 16-OCT-2001 (Rel. 40, Last annotation update)
DE Pulmonary surfactant-associated protein B precursor (SP-B) (Pulmonary
DE surfactant-associated proteolipid SPL(Phc)).
GN SFTPB OR SFTPB.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_Taxid=10090;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=DEA/20; TISSUE=Liver;
RX MEDLINE=95208782; Pubmed=7900819;
RA Bruno M.A., Bohinski R.J., Carter J.E., Fosse K.A., Wiltset J.A.;
RT "Structure and function of the mouse surfactant protein B gene.";
RL Am. J. Physiol. 268:1381-1389(1995).
CC -1- FUNCTION: PULMONARY SURFACTANT-ASSOCIATED PROTEINS PROMOTE
CC LIQUID INTERFACE BY LOWERING THE SURFACE TENSION AT THE AIR-
CC THE COLLAPESE PRESSURE OF PALMITIC ACID TO NEARLY 70 MILLINEWTONS
CC PER METER.
CC -1- SUBUNIT: Homodimer; disulfide-linked.
CC -1- SUBCELLULAR LOCATION: Extracellular.
CC -1- MISCELLANEOUS: Pulmonary surfactant consists of 90% lipid and 10%
CC protein. There are 4 surfactant-associated proteins: 2 collagenous,
CC carbohydrate-binding glycoproteins (SP-A and SP-D) and 2 small
CC hydrophobic proteins (SP-B and SP-C).
CC -1- SIMILARITY: Contains 1 saposin A-type domain.
CC -1- SIMILARITY: Contains 3 saposin B-type domains.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.1sb-sib.ch/announce/
CC or send an email to license@1sb-sib.ch).
CC -----
DR EMBL: S78114.1.AAB34846.2; -.
DR HSSP: P07988; 1DFW.
DR MGD: MGI:109516; Sftpb.
DR InterPro: IPR003119; Sapa.
DR InterPro: IPR007856; Sapa.1.
DR InterPro: IPR008138; Sapa.2.
DR InterPro: IPR008140; Sapa_sub.

```

```

DR InterPro: IPR008373; Saposin.
DR InterPro: IPR008139; SaposinB.
DR InterPro: IPR008137; Surfactant_B.
DR Pfam: PF02199; SAPA.1.
DR Pfam: PF05184; Sapa_1; 1.
DR Pfam: PF03489; Sapa_2; 3.
DR PRINTS: PR01797; SAPOSIN.
DR ProDom: PD001732; Sapa_sub; 1.
DR ProDom: PD006002; Surfactant_B; 1.
DR SMART: SM00162; SAPA; 1.
DR SMART: SM00118; Sapa; 3.
DR Surface Film; Gaseous exchange; Glycoprotein; Repeat.
FT CHAIN 1 191
FT PROPEP 192 270
FT FT 271 377 PULMONARY SURFACTANT-ASSOCIATED PROTEIN
FT DOMAIN 27 60 B.
FT DISULFID 199 268 SAPOSIN-LIKE TYPE A.
FT DISULFID 202 262 BY SIMILARITY.
FT DISULFID 226 237 BY SIMILARITY.
FT DISULFID 239 239 INTERCHAIN (BY SIMILARITY).
FT CARBOHYD 307 307 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 377 AA; 41728 MW; CB687A82BA3FC56C CRC64;

Query Match 13.7%; Score 378; DB 1; Length 377;
Best Local Similarity 22.5%; Pred. No. 9.5e-18;
Matches 111; Conservative 73; Mismatches 145; Indels 164; Gaps 16;

Dy 18 ECTGSAVMQCNKTKSDGAVHGCQTVMNKPVTSLPCDICKVYTAAGMLKDQATE 77
30 ECAQGPQFQMSIEHVAOCRAIGHCQEWGHAANDL-CQCEIDVHLITTKTEDEAFQ 88
Dy 78 EELVLEKTCMDLPFNNSASCKEIVDSYLVIIITIKGMSRPGVCSALNLCESLQK 137
89 EAIRKTELEOCEDLPFKLIVPRQGVLDVPLVIDYFQSLN-PKALCNHGLC----- 142
Dy 138 HLAELNHQQLSBNKIPBLDMTEVAVAFMANIPLLIYFQDGRSKPQPKNDGVCDCIQ 197
143 -----PRGAKPEON----- 152
Dy 198 MYTDIGTAVRTNSTFVQALVEHKECDRLGPGMADCKRYISGYSIAIQMAMHMQPK 257
153 -----PGMD----- 157
Dy 258 ICALVGFCEVEMEMQTLVPKASKNVTPALE-LVEP-----IKHEVPAKSDVY 308
158 -----AVPNPLLDKLVLPVLPALARPSPHTQDPSQQLPPLP-F 198
Dy 309 CEVCEFLVKEVTKLIDNNKTEKEILD-AFDKMSKLPKLSBEQOEVDVDTYGSSILSL 367
199 CWLCKRTLIKRVQAVI-----PKGVLAIVASQVCHVPLVVGICQCLAERTVLLDL 253
Dy 368 EIVSELYVCSMLHUCS-----GTRLPALTVAHTQ-PKDGFCVCEVKLVGLDRLNLEKNS 421
254 GRVNVQLVCGVILRSTEDMANGPALPAVEPLIEEPLQDTCHRCYSVI-----NQAMNT 308
Dy 422 TKQETIALKKGCC-SFLDDPYQKQCDQFVAEXEVEVLLEIVWVDPFVCLKTAGCSAH 480
309 SEQAMPQAMHQAARLRFWD-RQKCBQFVEQHMQLLALVRSQDAHITCOALGVCEAPA 366
Dy 481 KPLGTEKCIWGP 493
Dy 367 SPL-----QCFQTP 375

RESULT 8
PSPB_RAT STANDARD; PRT; 376 AA.
AC P22355;
DT 01-AUG-1991 (Rel. 19, Created)
DT 01-AUG-1991 (Rel. 19, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Pulmonary surfactant-associated protein B precursor (SP-B) (Pulmonary

```



DE surfactant-associated proteolipid SPL(Phe)).

GN SPTB OR SPT3.

OS Rattus norvegicus (Rat).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.  
OX NCBI\_TaxID=10116;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=89150284; PubMed=2920185;  
RA Emrie P.A., Shamon J.M., Mason R.J., Fisher J.H.,  
RT "cDNA and deduced amino acid sequence for the rat hydrophobic  
RT pulmonary surfactant-associated protein, SP-B,"  
RL Biochim. Biophys. Acta 994:215-221(1989).  
CC -! FUNCTION: PULMONARY SURFACTANT-ASSOCIATED PROTEINS PROMOTE  
CC ALVEOLAR STABILITY BY LOWERING THE SURFACE TENSION AT THE AIR-  
CC LIQUID INTERFACE IN THE PERIPHERAL AIR SPACES. SP-B INCREASES  
CC THE COLLAPSE PRESSURE OF PALMITIC ACID TO NEARLY 70 MILLINEWTONS  
CC PER METER.  
CC -! SUBUNIT: Homodimer; disulfide-linked.  
CC -! SUBCELLULAR LOCATION: Extracellular.  
CC -! MISCELLANEOUS: Pulmonary surfactant consists of 90% lipid and 10%  
CC protein. There are 4 surfactant-associated proteins: 2 collagenous,  
CC carbohydrate-binding glycoproteins (SP-A and SP-D) and 2 small  
CC hydrophobic proteins (SP-B and SP-C).  
CC -! SIMILARITY: Contains 1 saposin A-type domain.  
CC -! SIMILARITY: Contains 3 saposin B-type domains.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
DR EMBL; X14778; CAA32885.1; -  
DR PIR; S02766; S02766.  
DR HSSP; P07988; 1DFW.  
DR InterPro; IPR001119; Sapa.  
DR InterPro; IPR007856; Sapa\_1.  
DR InterPro; IPR008138; Sapa\_2.  
DR InterPro; IPR008140; Sapa\_sub.  
DR InterPro; IPR008373; Saposin.  
DR InterPro; IPR008139; SaposinB.  
DR InterPro; IPR008137; Surfactant\_B.  
DR Pfam; PF02119; Sapa\_1.  
DR Pfam; PF05184; Sapa\_1; 1.  
DR Pfam; PF03489; Sapa\_2; 3.  
DR PRINTS; PR01797; SAPOSIN.  
DR ProDom; PD001732; Sapa\_sub; 1.  
DR ProDom; PD008002; Surfactant\_B; 1.  
DR SMART; SM00162; SAPA; 1.  
DR SMART; SM00118; SAPB; 3.  
KW Surface film; Gaseous exchange; Glycoprotein; Repeat.  
FT PROPEP 1 190  
FT CHAIN 191 269 PULMONARY SURFACTANT-ASSOCIATED PROTEIN  
FT FT B.  
FT PROPEP 270 376  
FT DOMAIN 26 59 SAPOSIN-LIKE TYPE A.  
FT DISULFID 198 267 BY SIMILARITY.  
FT DISULFID 201 261 BY SIMILARITY.  
FT DISULFID 225 236 BY SIMILARITY.  
FT DISULFID 228 238 INTERCHAIN (BY SIMILARITY).  
FT CAROHID 306 306 N-LINKED (GLCNAC. . .) (POTENTIAL).  
SQ SEQUENCE 376 AA; 41590 MW; F329DC62E733FB4C CRC64;  
Query Match 13.5%; Score 372; DB 1; Length 376;  
Best Local Similarity 22.7%; Pred. No. 2.3e-17;  
Matches 115; Conservative 74; Mismatches 151; Indels 166; Gaps 17;  
QY 5 LGAALAGVLTGLKECTRGSAVWCQVKTASDCGAVHCLQVTWNNPTKYSLPFCIDCKDV 64  
DB 18 LGAAATES--ASSPDCAGPKFWCQSLQALICRALGHCLQVWGAANDL-CQCECDIV 74

QY 65 TAAGMLKDNATEBEILLVEKTCQDMLPKPMWSASCKEIVDSYLPVILDIKGEWSRPG 124  
DB 75 HLLRTMKEDAPFQDITRKFEQBCDILPLKLVPRCROVLVPLPLVDYFGQI-KKKA 133  
QY 125 VCSALNLCESLQKHLAEINHQKQLESNKIPELDMEVAVPPMANIPLLLYPQDGRSRKQ 184  
DB 134 ICSHVGLC-----PL---GQTRPQKPE 153  
QY 185 PKNDGVQDCQICQMTDITQIVRTNSFTVQALVEHVKECDRLGKADIKNTISQYSE 244  
DB 154 ----- 153  
QY 245 IAIQMMHMQKKEICALVGFCEVKEKMPQTLVPAKVAKNVIPALE---LVEP----- 295  
DB 154 ----- 185  
QY 296 IKHSEVPAKSDVYCEVECFVKEVYKLDNNKTEKILD-AFDKMSKTLPKSLSECOEV 354  
DB 186 LSEQQLPPLP-FCMLCKRTLKRQAVI-----PKGVAVAVSQCHVPLVVGICQCL 239  
QY 355 VDTGSSILSLILBEVSEPLVCSMLHCS-----GTRLPALTHTVQ-PKDGFCVCKK 408  
DB 240 AERYTVLLDLALGRVVPQVLCGLVRCSTADAIQPALPALLEPKWPLQDTCHFKCS 299  
QY 409 LVGYLDRLKNSRKQETLALKEGC-SFLDPYQKQDQVAAEYEPILILVEVMPDS 467  
DB 300 VI-----NQAWNTSQAMPQAMHQACRFWLD--RQKEQVEQHPOLLALVPRSDAH 352  
QY 468 FVCLKIGACPSAHKPLLTGTEKCIWCP 493  
DB 353 TSCQALGVCAPASPL-----QCFQTP 374

## RESULT 9

PSBP\_HUMAN STANDARD; PRT; 381 AA.  
ID PSBP\_HUMAN  
AC P07368; Q96R04;  
DT 01-AUG-1988 (Rel. 08, Created)  
DT 01-MAY-1992 (Rel. 22, Last sequence update)  
DT 10-OCT-2003 (Rel. 42, Last annotation update)  
DE Pulmonary surfactant-associated protein B precursor (SP-B) (6 kDa  
DE protein) (Pulmonary surfactant-associated proteolipid SPL(Phe)) (18  
DE kDa pulmonary-surfactant protein).  
GN SPTB OR SPT3.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=89170128; PubMed=2924687;  
RA Pilot-Matias T.J., Kister S.B., Fox J.L., Kropp K., Glaeser S.W.,  
RA Whitsett J.A.;  
RT "Structure and organization of the gene encoding human pulmonary  
RT surfactant proteolipid SP-B,"  
RL DNA 8:75-86(1989).  
RN [2]  
RP SEQUENCE FROM N.A., AND SEQUENCE OF 201-214.  
RC TISSUE=Lung;  
RX MEDLINE=87250653; PubMed=3597440;  
RA Jacobs K.A., Phelps D.S., Steindink R., Fisch J., Kriz R.,  
RA Jacobson L., Dougherty J.P., Raesch H.W., Flores J.;  
RT "Isolation of a cDNA clone encoding a high molecular weight precursor  
RT to a 6-kDa pulmonary surfactant-associated protein,"  
RL J. Biol. Chem. 262:9808-9811(1987).  
RN [3]  
RP SEQUENCE FROM N.A., AND VARIANTS ILE-131; PHE-176 AND HIS-272.  
RA Rieder M.J., Carlington D.P., Chung M.-W., Lee K.L., Poel C.L., Yi Q.,  
RA Nickerson D.A.;  
RL Submitted (JUL-2001) to the EMBL/Genbank/DBJ databases.  
RN [4]  
RP SEQUENCE FROM N.A.



Query Match 12.3%; Score 340.5; DB 1; Length 381;  
 Best Local Similarity 21.4%; Pred. No. 2.7e-15;  
 Matches 109; Conservative 71; Mismatches 169; Indels 161; Gaps 16;

QY 1 LASLIGALAGVIGLKECTRGSAVWCQNVKTASDCGAVKICLTQTVNKPPTVKSIPCDIC 60  
 DB 14 LPTLGGPTAAWTTSILACAGPEFWCQSLBQALQCRALGHCLQEWGHWGADDL-CQEC 72  
 QY 61 KDVAVNAADMDKDNATEBEILVLYLEKTCDMLEPKPMMSASCKEIVDSYLPVLIDIIKGENS 120  
 DB 73 EDIVHILNKKAKKALFQDTRKFLPQECNVLPFLKILMPQCNQVLDYFPLVIDYFQNDTD 132  
 QY 121 RPEGVASALNLCESLQKHLAEINHQKOLESNKI PELDMTEVAVAPPMANIPILLVYQDGR 180  
 DB 133 SNG-ICMHLGLCKSRQ----- 147  
 QY 181 SKPQKNDGVCQDCIQWVTDIQTAVRTNSTFVALVEHVKECDRLGPMADICKNYIS 240  
 DB 148 --PEPEQ-----PGMSD----- 158  
 QY 241 QYSEIAIQMMHMQKKEICALVGCDEVKEMQTLVPAKVASKNVLPALVEPIKKE 300  
 DB 159 -----PLPK----- 189  
 QY 301 VPAKSDV-----XCEVCEPLVKEVTKLIDNNKTEKEILDAPDKMSKLPKLSBE 350  
 DB 190 GPHODDLEEQEPPLPFLPCWLRALIKRIQAMIPRG---ALAVALVQCRVPLVAGGI 245  
 QY 351 COEVVDYTGSSILSLLEBSEBELVCSMLHLCS-----GTRLPALTVAHTQPKDGFCFV 405  
 DB 246 CQCLABERSVILDLTLGRMLPQLVCRILRCSMDSDAGPRSP---TGEWLPDRSE-CHL 301  
 QY 406 CCKLVGYIDRLNEKSTKQELIAALEKGC--SFLEDPYQKQCDQVAAEYEPVLIILEV 463  
 DB 302 CMSVT-----TQAGSSRQALPQAMIQACVGSMDL---REKCKQVEQHTPOLTLVPRG 353  
 QY 464 MDPFVCLIKIGACPSAHKPLLTGTEKICWGP 493  
 DB 354 WDAHTTCQALGVCGTMSPL-----QCINSP 379

RESULT 10  
 PSPB CANFA STANDARD; PRT; 363 AA.  
 AC P17129;  
 DT 01-AUG-1990 (Rel. 15, Created)  
 DT 01-AUG-1990 (Rel. 15, Last sequence update)  
 DT 16-OCT-2001 (Rel. 40, Last annotation update)  
 DE Pulmonary surfactant-associated protein B precursor (SP-B) (6 kDa protein) (Pulmonary surfactant-associated proteolipid SP-B) (Phe)  
 DE (Pulmonary surfactant protein 18) (SP 18) (Fragment).  
 GN SPTB OR SPTB3.  
 OS Canis familiaris (Dog).  
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.  
 OC NCBI\_Taxid=9615;  
 RN [1]  
 RP SEQUENCE FROM N.A., AND SEQUENCE OF 182-211.  
 RP TISSUE=Lung;  
 RX MEDLINE=87092398; PubMed=3467361;  
 RA Hawgood S., Benson B.J., Schilling J., Damm D., Clements J.A., White R.T.;  
 "Nucleotide and amino acid sequences of pulmonary surfactant protein SP 18 and evidence for cooperation between SP 18 and SP 28-36 in surfactant lipid adsorption.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 84:66-70 (1987).  
 CC -!- FUNCTION: PULMONARY SURFACTANT-ASSOCIATED PROTEINS PROMOTE ALVEOLAR STABILITY BY LOWERING THE SURFACE TENSION AT THE AIR-LIQUID INTERFACE IN THE PERIPHERAL AIR SPACES. SP-B INCREASES THE COLLAPSE PRESSURE OF PALMITIC ACID TO NEARLY 70 MILLINEWTONS PER METER.  
 CC -!- SUBUNIT: Homodimer; disulfide-linked.  
 CC -!- SUBCELLULAR LOCATION: Extracellular.

CC -!- MISCELLANEOUS: Pulmonary surfactant consists of 90% lipid and 10% protein. There are 4 surfactant-associated proteins: 2 collagenous, carbohydrate-binding glycoproteins (SP-A and SP-D) and 2 small hydrophobic proteins (SP-B and SP-C).  
 CC -!- SIMILARITY: Contains 1 saposin A-type domain.  
 CC -!- SIMILARITY: Contains 3 saposin B-type domains.  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation at the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see [http://www.isb-sib.ch/announce/or\\_send\\_an\\_email\\_to\\_license@sib.ch](http://www.isb-sib.ch/announce/or_send_an_email_to_license@sib.ch)).  
 CC EMBL, M15170; AAA30893.1; --  
 DR PIR, B29072; A29072.  
 DR HSSP, P07988; 1DFW.  
 DR InterPro: IPR003119; SABA.  
 DR InterPro: IPR007856; SABA\_1.  
 DR InterPro: IPR008138; SABA\_2.  
 DR InterPro: IPR008140; SABA\_sub.  
 DR InterPro: IPR008373; Saposin.  
 DR InterPro: IPR008137; SaposinB.  
 DR Pfam: PF02199; SABA\_1.  
 DR Pfam: PF05184; SABA\_1; 1.  
 DR Pfam: PF03489; SABA\_2; 3.  
 DR PRINTS, PR01797; Saposin.  
 DR ProDom: PD001732; SABA\_sub; 1.  
 DR ProDom: PD008002; Surfactant\_B; 1.  
 DR SMART, SM00162; SABA; 1.  
 DR SMART, SM00118; SABA; 3.  
 KW Surface film; Gaseous exchange; Glycoprotein; Repeat.  
 FT NON\_TER 1 180  
 FT PROPEP 1 259  
 FT CHAIN 181 259  
 FT PROPEP 260 363  
 FT DOMAIN 21 54  
 FT DISULFD 188 257  
 FT DISULFD 191 251  
 FT DISULFD 215 226  
 FT DISULFD 228 228  
 FT CARBOHYD 293 293  
 SQ SEQUENCE 363 AA; 40180 MW; PADDDQDDBB2719 CRC64;

Query Match 11.1%; Score 305; DB 1; Length 363;  
 Best Local Similarity 21.6%; Pred. No. 5.4e-13;  
 Matches 108; Conservative 75; Mismatches 159; Indels 158; Gaps 19;

QY 5 LGAA-LAGPVIGLKECTRGSAVWCQNVKTASDCGAVKICLTQTVNKPPTVKSIPCDIC 63  
 DB 13 LGADWMSAPSLA---CARGPAFWCQSLBQALQCRALGHCLQEWGHWGADDL-CQECQDI 68  
 QY 64 VTAAGDMKDNATEBEILVLYLEKTCDMLEPKPMMSASCKEIVDSYLPVLIDIIKGENS 123  
 DB 69 VRIITKTKKALFQDMWAKFLHEHEDVLPKILTPQCHNMGTFFPVVVDYFQSGIN-PK 127  
 QY 124 EVGSALNLCESLQKHLAEINHQKOLESNKI PELDMTEVAVAPPMANIPILLVYQDGR 183  
 DB 128 IICKHLGIC---KRLPEPEQESLSDPLKILPEL----- 162  
 QY 184 QPKQNDGVCQDCIQWVTDIQTAVRTNSTFVALVEHVKECDRLGPMADICKNYISQYS 243  
 DB 163 ----- 177  
 QY 244 EIALQMMHMQKKEICALVGCDEVKEMQTLVPAKVASKNVLPALVEPIKKEVPA 303  
 DB 178 E-----QQLP-----LP- 186  
 QY 304 KSDVYCEVCEPLVKEVTKLIDNNKTEKEILD-APDKMSKLPKLSSECCQEVVDYTGSSI 362

```

Db 187 -----YCWLCRTLIKRIQAMI-----PKGVLAIVTVGVGVHVVPLVVGICQCLGRRTVLL 237
Cc 363 LSILLEEVSPELVCSMLHLS-----GTRLPALVTVHTQPPDGGFCEVCKKLVGYLDRNL 417
Cc 238 LDALLGRLPQVGVGLVRCSHEDSAGPALSLPSE-WSIQESK-CQLCMFVT-----TQ 290
Cc 418 EKSTKQEIILALEKGC--SFLPDPYOKQCOQFAVEEPLYILVWMPSPVCLKIGA 475
Cc 291 AGNHSQEQTPQAIRQACLSLWLD---RQKCEQFVECHMPRLQTLASGRDHTTCQALGA 347
Cc 476 CPANHKLPTGTEKINMPSPY 495
Cc 348 CRTTFSP-----QCTHHPF 363

RESULT 11
PSPB RABIT
ID PSPB RABIT STANDARD; PRT; 370 AA.
AC P15285;
DT 01-APR-1990 (Rel. 14, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DE 16-OCT-2001 (Rel. 40, Last annotation update)
DE Pulmonary surfactant-associated protein B precursor (SP-B) (6 kDa protein) (Pulmonary surfactant-associated proteolipid SPL(phe)).
GN SFTPB OR SFTPB3
OS Oryctolagus cuniculus (Rabbit).
OC Eukaryota; Metazoa; Chordata; Cranata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.
OX NCBI_TaxID=9986;
RN 1
RP SEQUENCE FROM N.A.
RC MEDLINE=89228033; PubMed=2469419;
RX TISSUE=Lung;
RA Xu J., Richardson C., Ford C., Spencer T., Li-Juan Y., Mackie G.,
RT Hammond G., Possmayer F.;
RT "Isolation and characterization of the cDNA for pulmonary surfactant-associated protein-B (SP-B) in the rabbit."
RL Biochem. Biophys. Res. Commun. 160:325-332(1989).
RN 12
RP SEQUENCE FROM N.A.
RC STRAIN=New Zealand white;
RX MEDLINE=95208794; PubMed=7900830;
RA Maxgana R.K., Boggaram V.;
RT "Transcription and mRNA stability regulate developmental and hormonal expression of rabbit surfactant protein B gene."
RL Am. J. Physiol. 268:1481-1490(1995).
RN 13
RP SEQUENCE OF 1-34 FROM N.A.
RX MEDLINE=96096536; PubMed=8522191;
RA Lutz P., Anceleschi M., Strayer D.S.;
RT "The upstream region of the SP-B gene: intrinsic promoter activity and glucocorticoid responsiveness related to a new DNA-binding protein."
RL Gene 165:285-290(1995).
Cc -1- FUNCTION: PULMONARY SURFACTANT-ASSOCIATED PROTEINS PROMOTE LIQUID STABILITY BY LOWERING THE SURFACE TENSION AT THE AIR-LIQUID INTERFACE IN THE PERIPHERAL AIR SPACES. SP-B INCREASES THE COLLAPSE PRESSURE OF PALMITIC ACID TO NEARLY 70 MILLINEWTONS PER METER.
Cc -1- SUBUNIT: Homodimer; disulfide-linked.
Cc -1- SUBCELLULAR LOCATION: Extracellular.
Cc -1- MISCELLANEOUS: Pulmonary surfactant consists of 90% lipid and 10% protein. There are 4 surfactant-associated proteins: 2 collagenous, carbohydrate-binding glycoproteins (SP-A and SP-D) and 2 small hydrophobic proteins (SP-B and SP-C).
Cc -1- SIMILARITY: Contains 1 saposin A-type domain.
Cc -1- SIMILARITY: Contains 3 saposin B-type domains.
Cc -----
Cc This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation at the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial

```

```

Cc entities requires a license agreement (See http://www.isb-sib.ch/announce/
Cc or send an email to license@isb-sib.ch).
Cc -----
Cc EMBL; M24901; AAA31466.1; -
Cc EMBL; U17106; AA67934.1; -
Cc EMBL; S80649; AADI435.1; -
Cc PIR; A32421; LNRBB.
Cc PIR; I46531; I46531.
Cc HSSP; P07988; IDPW.
Cc InterPro; IPR003119; Sapa.
Cc InterPro; IPR007856; Sapa_1.
Cc InterPro; IPR008138; Sapa_2.
Cc InterPro; IPR008140; Sapa_sub.
Cc InterPro; IPR008139; SaposinB.
Cc InterPro; IPR008137; Surfactant_B.
Cc Pfam; PF02199; SABA; 1.
Cc Pfam; PF03489; Sapa_1; 1.
Cc Pfam; PF03489; Sapa_2; 3.
Cc ProDom; PD001732; Sapa_sub; 1.
Cc ProDom; PD008002; Surfactant_B; 1.
Cc SMART; SM00162; SABA; 1.
Cc SMART; SM00118; SABA; 3.
Cc Surface film; Gaseous exchange; Glycoprotein; Repeat.
Cc PROPEP
Cc CHAIN
Cc 185 263
Cc 184 263
Cc 184 263
Cc PROPEP
Cc 264 370
Cc 29 62
Cc DOMAIN
Cc 192 261
Cc DISULFID
Cc 195 255
Cc DISULFID
Cc 219 230
Cc DISULFID
Cc 232 232
Cc CARBOHYD
Cc 300 300
Cc CONFLICT
Cc 15 15
Cc CONFLICT
Cc 184 184
Cc CONFLICT
Cc 232 232
Cc CONFLICT
Cc 329 355
Cc SEQUENCE 370 AA; 40610 MW; 423047A69B12DCB5 CRC64;
Cc -----
Cc Query Match 10.5%; Score 290; DB 1; Length 370;
Cc Best Local Similarity 21.6%; Pred. No. 5.3e-12;
Cc Matches 111; Conservative 65; Mismatches 167; Indels 170; Gaps 19;
Cc -----
Cc 1 LSILGALAGLVLDL-----KECTGSAVWCQVKTASDCAVYKCLQTVWKNPKYKSL 55
Cc 10 LILLILPTLCGFTAVWATSPPLACQGEFECOSILEQALCKALGHCLQEWGHHVADDL 69
Cc 56 PCDICKDVVTAAGMLKDNATEBEILVLEKTCQMLPKPNMSASCKEILVSYPLVILDI 115
Cc 70 -QCEQDILVILTKMTKALFQDITRKLEHCEVLPKILVPCQHHVLDVFPPLTTFP 128
Cc 116 KQEMRBEVCSALNLCESLQKHLELNHOKLESNKIPILDMTEVVAAPMANIPLLLYP 175
Cc 129 OSQINAKA-ICQHGGLQ-----PGSDEPLD-----PLPKVLVPTLL- 166
Cc 176 QDGPSEKQPPQNDGVQDCIQWTVTDIQAVRNSTSTVQVLYVHVEKSCRLPGNADIC 235
Cc 167 -GALPARK-----SPHTDLS 181
Cc 236 KNYISQISEAIQMMHMPKEICALVGFDEVEMQMQLVPAKVASKNVLPALVELVE 295
Cc 182 -----AQRPIPL-----P 190
Cc 296 IKKEHVAKSDVVGCEVEFLVKEVTKLIDNNKTEKEILD-APDKMSKLPKSLSESCQEV 354
Cc 191 L-----CWLCTLIKRIQAMI-----PKGVLAIVTVGVGVHVVPLVVGICQCL 233
Cc 355 VDTYGSILSILLEEVSPELVCSMLHLS-----GTRLPALVTVHTQ--PKDGGFCEVCK 407
Cc 234 AERYTVILLVLLGVHLPQVGVGLVRCSSVDSIGVPPVLELPEGEMLPDDE-CRLDM 292
Cc 408 KLVGYLDRNLKSTKQEIILAA-----LEKGSFLPDPYOKQCOQFAVEEPLYILVIVE 462

```

```
Db 293 SVTQA-RNISEQRPAQVAVHACUSQLDK-----GECEQFVELHPTQLSLSR 341
Qy 463 VMDPSFVLKLGACSAKPLIGTEKTCWGPST 495
Db 342 GMDARAIQALGACVATLSPF-----QCTIQSPHF 370

RESULT 12
SAP_CAVPO STANDARD; PRT; 81 AA.
ID SAP_CAVPO STANDARD; PRT; 81 AA.
AC P20097;
DT 01-FEB-1991 (Rel. 17, Created)
DT 01-FEB-1991 (Rel. 17, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Saposin C (CO-beta-glucosidase) [Glucosylceramidase activator]
DE (Sphingolipid activator protein 2) (SAP-2).
OS Cavia porcellus (Guinea pig).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Hystricognathi; Caviidae; Cavia.
OX NCBI_TaxID=10141;
RN [1]
RP SEQUENCE.
RC TISSUE=Liver;
RX MEDLINE=89066787; PubMed=3198642;
RA Sano A., Radin N.S., Johnson L.L., Tarr G.E.;
RT "The activator protein for glucosylceramide beta-glucosidase from
RT guinea pig liver. Improved isolation method and complete amino acid
RT sequence."
RL J. Biol. Chem. 263:19597-19601(1988).
CC -1- FUNCTION: Saposin A and saposin C stimulate the hydrolysis of
CC glucosylceramide by beta-glucosylceramidase (EC 3.2.1.45) and
CC galactosylceramide by beta-galactosylceramidase (EC 3.2.1.46).
CC Saposin C apparently acts by combining with the enzyme and acidic
CC lipid to form an activated complex, rather than by solubilizing
CC the substrate.
CC -1- SIMILARITY: Contains 1 saposin B-type domain.
CC PIR: A32026; A32026.
DR HSSP; P42210; 10DM.
DR InterPro; IPR007856; SapB_1.
DR InterPro; IPR008138; SapB_2.
DR InterPro; IPR008140; SapB_sub.
DR InterPro; IPR008173; Saposin.
DR InterPro; IPR008139; SaposinB.
DR Pfam; PF05184; SapB_1; 1.
DR Pfam; PF05184; SapB_2; 1.
DR Pfam; PF03489; SapB_2; 1.
DR PRINTS; PR01797; SAPOSIN.
DR PRODOM; PD001732; SapB_sub; 1.
DR SMART; SM00118; SAPB; 1.
DR KW Glycoprotein; Sphingolipid metabolism.
FT DOMAIN 1 81 SAPOSIN-LIKE TYPE B.
FT CARBOHD 22 22 N-LINKED (GLCNAC. . .).
FT DISULFD 5 78 BY SIMILARITY.
FT DISULFD 8 72 BY SIMILARITY.
FT DISULFD 36 47 BY SIMILARITY.
FT SEQUENCE 81 AA; 8852 MW; E564CE1F0A292556 CRC64;

Query Match 9.5%; Score 261; DB 1; Length 81;
Best Local Similarity 61.8%; Pred. No. 6.4e-11;
Matches 47; Conservative 18; Mismatches 11; Indels 0; Gaps 0;

Qy 307 VYCEVCELVKVKTLINNTKEKILDAFKMCKSLPKSLSECEQVVDYTGSSILSL 366
Db 3 VTCACACEVAVKVKVELLINNTKEKITHALDSVCLLPESVSEVCEQVVDYTGSSIVALL 62

Qy 367 LEEVSPELVCSMLHLC 382
Db 63 LQEMSPELVCSMLHLC 78

RESULT 13
ASPR_CUCPE STANDARD; PRT; 513 AA.
ID ASPR_CUCPE STANDARD; PRT; 513 AA.
```

```
AC 004057;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Aspartic proteinase precursor (EC 3.4.23.-).
OS Cucurbita pepo (Vegetable marrow) (Summer squash).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC euroside I; Cucurbitales; Cucurbitaceae; Cucurbita.
OX NCBI_TaxID=3663;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=97354181; PubMed=9210475;
RA Hiraiwa N., Kondo M., Nishimura M., Hara-Nishimura I.;
RT "An aspartic endopeptidase is involved in the breakdown of
RT propeptides of storage proteins in protein-storage vacuoles of
RT plants."
RL Eur. J. Biochem. 246:133-141(1997).
CC -1- FUNCTION: Involved in the breakdown of propeptides of storage
CC proteins in protein-storage vacuoles.
CC -1- SUBUNIT: Heterodimer of a 32 kDa subunit and a 16 kDa subunit.
CC -1- SUBCELLULAR LOCATION: Vacuolar.
CC -1- SIMILARITY: Belongs to peptidase family A1.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AB002695; BAA19607.1; -.
DR PIR; T09739; T09739.
DR HSSP; P42210; 10DM.
DR MEROPS; A01.020; -.
DR InterPro; IPR001969; Asprotease_AS.
DR InterPro; IPR009007; Pept_A_acid.
DR InterPro; IPR001461; Peptidase_A1.
DR InterPro; IPR007856; SapB_1.
DR InterPro; IPR008138; SapB_2.
DR InterPro; IPR008140; SapB_sub.
DR InterPro; IPR008173; Saposin.
DR InterPro; IPR008139; SaposinB.
DR Pfam; PF00026; asp; 1.
DR Pfam; PF05184; SapB_1; 1.
DR Pfam; PF03489; SapB_2; 1.
DR PRINTS; PR00792; PERSIN.
DR PRINTS; PR01797; SAPOSIN.
DR PRODOM; PD001732; SapB_sub; 1.
DR SMART; SM00118; SAPB; 2.
DR PROSITE; PS00141; ASP_PROTEASE; 2.
DR KW Hydroxylase; Aspartyl protease; Zymogen; Glycoprotein; Signal.
FT SIGNAL 1 25 POTENTIAL.
FT PROPEP 26 513 POTENTIAL.
FT CHAIN 26 513 ASPARTIC PROTEINASE.
FT DOMAIN 322 420 SPECIFIC TO PLANT ASPARTIC PROTEINASES.
FT ACT_SITE 108 108 BY SIMILARITY.
FT ACT_SITE 294 294 BY SIMILARITY.
FT DISULFD 120 126 BY SIMILARITY.
FT DISULFD 285 289 BY SIMILARITY.
FT DISULFD 432 469 BY SIMILARITY.
FT CARBOHD 143 143 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHD 404 404 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT SEQUENCE 513 AA; 55855 MW; 083FB7064CE02DC2 CRC64;

Query Match 6.3%; Score 173.5; DB 1; Length 513;
Best Local Similarity 29.0%; Pred. No. 0.00034;
Matches 47; Conservative 31; Mismatches 59; Indels 25; Gaps 7;

Qy 347 LSECEQVVDYTGSSILSLLEEVSPELVCSMLHCS--GTRLPALTVH-VTQPKDG--- 400
Db 320 VSQCKAVVAQYGGTINDLLSEADPKKICQINLCTFDGRGVSGVIBSVVDENAGKSS 379
```

Oy 401 -----GFCVCKLVGLDRMLKNSKQELALKEGCSFLPPYOK---QCDQPAEY 452  
 Db 380 DSLHDMGVCSEVMVMMQNLRONQKTERINININECDRMEPMQSAVDGQ-LSSM 438  
 Oy 453 EPLYELIVEVD--PSFVCKIKACPSANPKPLGTEKCIWG 492  
 Db 439 PTVSFTIGKIFDLAPERYILKVEGCPVA-----QCISG 472

## RESULT 14

ASPR\_HORVU ID ASPR\_HORVU STANDARD, PRT, 508 AA.

AC P42210; 01-NOV-1995 (Rel. 32, Created)

DT 01-NOV-1995 (Rel. 32, Last sequence update)

DE 15-MAR-2004 (Rel. 43, Last annotation update)

OS Phytasein precursor (EC 3.4.23.40) (Aspartic proteinase).

OC Hordeum vulgare (Barley).

OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooidae;

OC Triticeae; Hordeum.

NCBI\_TaxID=4513;

[1] SEQUENCE FROM N.A.

RC STRAIN=cv. Kustaa; TISSUE=Embryo;

RA MEDLINE=92111473; PubMed=1722454;

RT Rumberg-Roos P., Toermaekangas K., Oestman A.;

RL "Primary structure of a barley-grain aspartic proteinase. A plant

aspartic proteinase resembling mammalian cathepsin D.";

Eur. J. Biochem. 202:1021-1027(1991).

[2] SEQUENCE FROM N.A. AND PARTIAL SEQUENCE.

RC STRAIN=cv. Kustaa; TISSUE=Embryo;

RA MEDLINE=92254717; PubMed=1812727;

RT Toermaekangas K., Rumberg-Roos P., Oestman A., Tilgmann C.,

Sarkkinen P., Kervinen J., Mikola I., Kalkkinen N.;

"Aspartic proteinase from barley seeds is related to animal cathepsin

D.";

Adv. Exp. Med. Biol. 306:355-359(1991).

[3] X-RAY CRYSTALLOGRAPHY (2.3 ANGSTROMS).

RA MEDLINE=9335466; PubMed=10406799;

RT Kervinen J., Tobin G.J., Costa J., Maugh D.S., Wlodawer A., Zdanov A.;

"Crystal structure of plant aspartic proteinase prophytepsin:

inactivation and vacuolar targeting.";

EMBO J. 18:3947-3955(1999).

-1- CATALYTIC ACTIVITY: Prefers hydrophobic residues Phe, Val, Ile,

Leu, and Ala at P1 and P1', but also cleaves -Phe-|-Asp- and

-Asp-|-Asp- bonds in 2S albumin from plants seeds.

-1- SUBUNIT: Heterodimer of two subunits (29 kDa and 11 kDa) processed

from the precursor molecule. A large enzyme (32 kDa and 16 kDa) is

an immediate precursor form.

-1- SUBCELLULAR LOCATION: Vacuolar.

-1- TISSUE SPECIFICITY: Embryo and leaf.

-1- SIMILARITY: Belongs to peptidase family A1.

-----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.1sb-sib.ch/announce/>  
 CC or send an email to [license@1sb-sib.ch](mailto:license@1sb-sib.ch)).

EMBL: X56136, CAA39602.1, -

PIR: S19697, S19697.

PDB: 1QDM, 16-JUL-99.

MEROPS: A01.020.1.

InterPro: IPR001969; Asparticase AS.

InterPro: IPR009007; Pept A acid.

InterPro: IPR001461; Peptidase A1.

InterPro: IPR007856; Sapp\_1.

DR InterPro: IPR008138; Sapp\_2.  
 DR InterPro: IPR008140; Sapp\_sub.  
 DR InterPro: IPR008373; Saposin.  
 DR InterPro: IPR008139; SaposinB.  
 DR Pfam: PF00026; aap; 1.  
 DR Pfam: PF05184; Sapp\_1; 1.  
 DR Pfam: PF03489; Sapp\_2; 1.  
 DR PRINTS: PR00792; PEPSIN.  
 DR PRINTS: PR01797; SAPOSIN.  
 DR ProDom: PD001732; Sapp\_sub, 1.  
 DR SMART: SM00118; Sapp\_2.  
 DR PROSITE: PS00141; ASP\_PROTEASE, 2.  
 DR Hydrolase; Aspartyl protease; Zymogen; Glycoprotein; Signal;  
 KM 3D-structure.  
 KW SIGNAL 1 22  
 FT PROPEP 23 66  
 FT CHAIN 67 377  
 FT CHAIN 67 377  
 FT CHAIN 67 377  
 FT CHAIN 378 508  
 FT CHAIN 422 508  
 FT DOMAIN 317 420  
 FT SITE 377 378  
 FT SITE 421 422  
 FT ACT\_SITE 102 102  
 FT ACT\_SITE 289 289  
 FT DISULFID 115 121  
 FT DISULFID 280 284  
 FT DISULFID 427 464  
 FT CARBOHYD 399 399  
 FT STRAND 32 37  
 FT HELIX 42 49  
 FT HELIX 50 50  
 FT HELIX 74 76  
 FT HELIX 79 81  
 FT STRAND 84 90  
 FT STRAND 91 94  
 FT STRAND 95 102  
 FT STRAND 103 104  
 FT STRAND 109 112  
 FT TURN 113 114  
 FT HELIX 119 123  
 FT STRAND 127 127  
 FT HELIX 129 131  
 FT TURN 133 134  
 FT STRAND 136 136  
 FT STRAND 141 146  
 FT TURN 147 148  
 FT STRAND 149 162  
 FT TURN 163 164  
 FT STRAND 165 177  
 FT STRAND 180 180  
 FT HELIX 182 185  
 FT STRAND 190 193  
 FT HELIX 197 199  
 FT HELIX 201 203  
 FT HELIX 207 211  
 FT TURN 212 215  
 FT STRAND 221 225  
 FT STRAND 236 240  
 FT TURN 241 241  
 FT TURN 245 246  
 FT STRAND 248 260  
 FT TURN 261 262  
 FT STRAND 263 267  
 FT STRAND 270 272  
 FT TURN 273 274  
 FT STRAND 275 276  
 FT TURN 279 282  
 FT STRAND 284 288  
 FT STRAND 295 297  
 FT HELIX 299 309  
 FT TURN 310 310  
 FT STRAND 314 314

N-LINKED (GLCNAC. . .) (POTENTIAL).

POTENTIAL.  
 PHYTERPIN 32 kDa SUBUNIT.  
 PHYTERPIN 29 kDa SUBUNIT.  
 PHYTERPIN 16 kDa SUBUNIT.  
 PHYTERPIN 11 kDa SUBUNIT.  
 SPECIFIC TO PLANT ASPARTIC PROTEINASES.  
 CLEAVAGE.

FT HELIX 317 325  
 FT TURN 326 326  
 FT HELIX 327 335  
 FT TURN 336 337  
 FT HELIX 340 346  
 FT TURN 347 348  
 FT HELIX 379 397  
 FT TURN 398 399  
 FT HELIX 402 412  
 FT TURN 413 414  
 FT STRAND 416 416  
 FT STRAND 424 425  
 FT HELIX 427 430  
 FT TURN 431 432  
 FT STRAND 436 440  
 FT TURN 441 442  
 FT STRAND 443 447  
 FT HELIX 449 452  
 FT STRAND 453 455  
 FT HELIX 459 461  
 FT STRAND 464 466  
 FT TURN 468 470  
 FT STRAND 475 476  
 FT STRAND 480 483  
 FT HELIX 485 488  
 FT TURN 489 490  
 FT STRAND 491 496  
 FT TURN 497 500  
 FT STRAND 501 508  
 SQ SEQUENCE 508 AA; 54226 MW; 87F2C9F93369B962 CRC64;

Query Match 5.7%; Score 158.5; DB 1; Length 508;  
 Best Local Similarity 24.3%; Pred. No. 0.0032;  
 Matches 49; Conservative 40; Mismatches 78; Indels 35; Gaps 9;

187 DNGDV-----CODCIOWTDTOTAVRTSTFVALVHEVKECDRLGPMADICN 237  
 DB DMGDLVVGKSTGFCAGCAALADSGTSLAGPT---AITEINKIGAGV-VEQCKT 321  
 QY 238 YISQYSEIAIQMM-HMQPKICALVGFCDVEKEMPMQTLVPAKVASKNVIPALEVPT 296  
 DB 322 IVSGYGOQILDLALAQTPKIKCSQVGLC-----TPDGTGVSAGIRSVVD-DEPV 371  
 QY 297 KKHVPASDVYCECFELVKEVTILDNKTEKILDAFDMGCKLTKSLSE--EC-- 351  
 DB 372 KSNGL--RADPVCASACMAVVMQNLQNKTDLLIVYQLCNRLSPMGESAVDCGS 429  
 QY 352 ----QEVVDYGGSSILILEE 369  
 DB 430 LGSMPDIEFTIGCKKFKALPKEE 451

RESULT 15  
 CYP1\_CYNCA STANDARD; PRT; 473 AA.  
 AC P40782:  
 DT 01-FEB-1995 (Rel. 31, Created)  
 DT 01-NOV-1995 (Rel. 32, Last sequence update)  
 DT 28-FEB-2003 (Rel. 41, Last annotation update)  
 DE Cyprosin precursor (EC 3.4.23.-) (Fragment).  
 GN CYP1.  
 OS Cynara cardunculus (Cardoon).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; asterids;  
 OC Campanulids; Asterales; Asteraceae; Carduoideae; Cardueae; Cynara.  
 OX NCBI\_TaxID=4265;  
 RN [1]  
 RP SEQUENCE FROM N.A., AND SEQUENCE OF 178-186.  
 RC TISSUE=Flower buds;  
 RX MEDLINE=94250836; PubMed=8193298;  
 RX Cordeiro M.C., Xue Z.-T., Pietrzak M., Pais M.S., Brodelius P.E.,  
 RT "Isolation and characterization of a cDNA from flowers of Cynara  
 cardunculus encoding cyprosin (an aspartic proteinase) and its use to

RT study the organ-specific expression of cyprosin.";  
 RL Plant Mol. Biol. 24:733-741(1994).  
 CC -1- TISSUE SPECIFICITY: Mostly present in the violet parts of styles  
 CC and corollas of mature flowers.  
 CC -1- DEVELOPMENTAL STAGE: Expressed in early stages of floral  
 CC development and switched off at maturation of the flower.  
 CC -1- SIMILARITY: Belongs to peptidase family A1.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL; X69193; CAA48939.1; ALT\_SEQ.  
 DR HSSP; P42210; IQDM.  
 DR MEROPS; A01.020; -.  
 DR InterPro; IPR001969; Asparticase\_AS.  
 DR InterPro; IPR009007; Pept\_A\_acid.  
 DR InterPro; IPR001461; Peptidase\_A1.  
 DR InterPro; IPR007856; SAb 1.  
 DR InterPro; IPR008138; SAb 2.  
 DR InterPro; IPR008140; SAb sub.  
 DR InterPro; IPR008139; SapostinB.  
 DR Pfam; PF00026; asp; 1.  
 DR Pfam; PF05184; SAb 1; 1.  
 DR Pfam; PF03489; SAb 2; 1.  
 DR PRINTS; PR00792; PERSIN.  
 DR ProDom; PD001732; SAb sub; 1.  
 DR SMART; SM00118; SAb; 2.  
 DR PROSITE; PS00141; ASP\_PROTEASE; 2.  
 DR KMW Hydroxylase; Aspartyl protease; Zymogen; Glycoprotein.

FT NON TER 1 1  
 FT PROBEP <1 33  
 FT CHAIN 34 473  
 FT ACT\_SITE 69 256  
 FT ACT\_SITE 256 256  
 FT DISULFID 82 88  
 FT DISULFID 247 251  
 FT DISULFID 392 429  
 FT CARBOHYD 364 364  
 FT CARBOHYD 411 411  
 SQ SEQUENCE 473 AA; 51564 MW; 65F3232BED06CB56 CRC64;

Query Match 5.6%; Score 154.5; DB 1; Length 473;  
 Best Local Similarity 30.1%; Pred. No. 0.0054;  
 Matches 31; Conservative 24; Mismatches 39; Indels 9; Gaps 2;

QY 347 LSEEGQEVVDYGGSSILILEVSPBLVGMHLCS--GTRLPALTVHTQPKDGS--- 401  
 DB 282 MSQOCKSLVDQYDGKSMIEMLSEEPKICSGMKLCSFDGSHDTSMTIESVVDKSKGKS 341  
 QY 402 ----FCEVCKTLVGLDNLNLEKNSKTOETLLALEKGSFLPDP 440  
 DB 342 GLPMKCYPCARVVMQNLQNKTEENTINYVDLCLRLPSP 384

Search completed: May 5, 2004, 13:31:43  
 Job time : 13.7901 secs

**THIS PAGE BLANK (USPTO)**



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 5, 2004, 13:25:43 ; Search time 51.6524 Seconds  
(without alignment)  
3164.197 Million cell updates/sec

Title: US-09-743-684a-1\_COPY\_7\_524  
Perfect score: 2759  
Sequence: 1 LASLGAALAGPVGLKECT.....NTETAQCNVAVHCRRHYWN 518

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 1017041 seqs, 315518202 residues  
Total number of hits satisfying chosen parameters: 1017041

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Listing first 45 summaries

Database :

1: SPREMBL\_25:\*  
2: sp\_archaea:\*  
3: sp\_bacteria:\*  
4: sp\_fungi:\*  
5: sp\_human:\*  
6: sp\_invertebrate:\*  
7: sp\_mammal:\*  
8: sp\_mhc:\*  
9: sp\_organelle:\*  
10: sp\_plant:\*  
11: sp\_todent:\*  
12: sp\_virus:\*  
13: sp\_vertebrate:\*  
14: sp\_unclassified:\*  
15: sp\_virus:\*  
16: sp\_bacteriap:\*  
17: sp\_archaeap:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1928	69.9	554	11	Q8BFQ1 mus musculus
2	1437	52.1	512	13	Q7BY70 xenopus lae
3	1412	51.2	520	13	Q8UVZ4 mus musculus
4	1341	48.6	522	13	Q9DG82 brachydanio
5	1089.5	39.5	525	11	Q8C1C1 mus musculus
6	776	28.1	449	11	Q8B7V5 mus musculus
7	772	28.0	402	11	Q8C1N0 mus musculus
8	549.5	19.9	953	5	Q9Y125 drosophila
9	545	19.8	241	4	Q8N7T4 homo sapien
10	517	18.7	965	6	P79254 ovis aries
11	511	18.5	921	5	O15997 bombyx mori
12	432.5	15.7	876	5	Q8IMH4 drosophila
13	343.5	12.5	458	5	Q95X02 naegleria f
14	343.5	12.4	484	5	Q9BMM1 naegleria f
15	342.5	12.4	378	11	O35489 cavia porce
16	330.5	12.0	374	6	Q9TUB1 ovis aries

17	327.5	11.9	441	5	Q9UPP3 drosophila
18	288	10.4	307	5	Q9BKM2 naegleria f
19	286	10.4	370	6	P79333 oryctolagus
20	279	10.1	294	5	Q95X03 naegleria f
21	279	10.1	456	5	Q94472 dictyosteli
22	251	9.1	316	5	Q86PA4 drosophila
23	220.5	8.0	429	5	O18276 caenorhabdi
24	213	7.7	213	10	Q9SC75 arabidopsis
25	203.5	7.4	217	10	Q9LZW6 arabidopsis
26	187.5	6.8	243	6	Q9T705 ovis aries
27	180	6.5	200	5	Q86JDB dictyosteli
28	179.5	6.5	240	10	Q9AS89 oryza sativ
29	176.5	6.4	402	5	O18279 caenorhabdi
30	176	6.4	228	6	Q9BDZ9 ovis aries
31	173.5	6.3	174	10	Q9M614 vitis ripar
32	173.5	6.3	507	10	Q9FRK7 nepenthes a
33	172	6.2	188	5	Q86KA8 dictyosteli
34	171.5	6.2	486	10	Q38934 arabidopsis
35	171.5	6.2	506	10	O65390 arabidopsis
36	165.5	6.0	506	10	Q39311 brassica na
37	163.5	5.9	273	10	Q9LUX5 pyrus pyrif
38	163.5	5.9	513	10	Q8VY13 arabidopsis
39	162.5	5.9	513	10	Q41713 vigna ungu
40	162	5.9	514	10	Q8L6A9 theobroma c
41	161	5.8	107	10	Q8H0R9 cynara card
42	157	5.7	512	10	O04593 arabidopsis
43	156.5	5.7	514	10	Q941A2 glycine max
44	155.5	5.6	205	6	Q9N275 ovis aries
45	155.5	5.6	292	10	Q43407 brassica o1

#### ALIGNMENTS

RESULT 1  
Q8BFQ1 PRELIMINARY; PRT; 554 AA.  
ID Q8BFQ1  
AC Q8BFQ1  
DT 01-MAR-2003 (TREMBLrel. 23, Created)  
DT 01-MAR-2003 (TREMBLrel. 23, Last sequence update)  
DT 01-OCT-2003 (TREMBLrel. 25, Last annotation update)  
DE Prosaposin.  
GN PSAP.  
OS Mus musculus (Mouse).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
OX NCBI\_TaxID=10090;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=NOB; TISSUE=Kidney, and Thymus;  
RC MEDLINE=22354683; PubMed=12466851;  
RA The FANTOM Consortium,  
RA the RIKEN Genome Exploration Research Group Phase I & II Team;  
RT "Analysis of the mouse transcriptome based on functional annotation of  
60,770 full-length cDNAs."  
RL Nature 420:563-573(2002).  
DR EMBL; AK086369; BAC40308.1; -  
DR EMBL; AK089998; BAC41035.1; -  
DR MGD; MGI:97783; Psap.  
DR GO; GO:0005764; C:lysosome; IEA.  
DR GO; GO:0006665; P:sphingolipid metabolism; IEA.  
DR InterPro; IPR003119; SAPA.  
DR InterPro; IPR007856; SAPB\_1.  
DR InterPro; IPR008138; SAPB\_2.  
DR InterPro; IPR008140; SAPB\_sub.  
DR InterPro; IPR008373; Saposin.  
DR InterPro; IPR008139; SaposinB.  
DR Pfam; PF02199; SAPA; 2.  
DR Pfam; PF05184; SAPB\_1; 4.  
DR Pfam; PF03489; SAPB\_2; 4.  
DR PRINTS; PR01797; SAPOSIN.  
DR ProDom; PD001732; SAPB\_sub; 3.  
DR SMART; SM00162; SAPA; 2.

DR SMART; SMO0118; SAMP; 4.  
SQ SEQUENCE 554 AA; 61050 MW; FF58DB79C7CC0C18 CRC64;  
Query Match 69.9%; Score 1928; DB 11; Length 554;  
Best Local Similarity 63.9%; Pred. No. 3e-138;  
Matches 350; Conservative 78; Mismatches 88; Indels 32; Gaps 2;  
QY 2 ASLIGALAGPVLGKCTRGSAVWCONVKTASDCGAVHGCLOTWNNKPVVSLPDDICK 61  
DB 8 ASLIALATLSPVDPDKTCSGSAVLCDVITADCGAVHGCQOMVMSKPLASLPDDICK 67  
QY 62 DVTTAGADMKNATBEELIVLEKTCMDLPRKNMSASCKEIVDSYLPYILDIKGMER 121  
DB 68 TVTTEAGNMLKDAVTEELIHLTEKTCMEIHDSSLASCEVDSYLPYILDMIGEMN 127  
QY 122 PGEVCSALNLCESLQKHAELNHOKEKSKIPELDMTEVAVPAPNITLILYPODGERS 181  
DB 128 PGEVCSALNLCESLQKHAELNHOKEKSKIPELDMTEVAVPAPNITLILYPODGERS 186  
QY 182 KPQKXNDGVQDCIQWVTDIQTAVRTNSTFVQALVEHVEKEDCDRLGPGMADICKNYISQ 241  
DB 187 QPQKXNDGVQDCIQWVTDIQTAVRTNSTFVQALVEHVEKEDCDRLGPGMADICKNYISQ 246  
QY 242 YSEIALOMMMHOPKEICLVGFCDEVKEMQTLVPAKVASKNVLPALTEVPIKHEV 301  
DB 247 YSEVVCVQMLHMOPEKICVLGACNEVKVPMKTLVPAETIKNILLPELMMDPYEQNLV 306  
QY 302 PAKSDVYCEVCEFLVETKLIIDNNKTEKILIDADPKMSCKPKSLSECCQEVDTYSS 361  
DB 307 QAHNVILCQTCQFVNMKSEELIVNATBELVAKGLSNACALLPDPARTKCEVGTGFP 366  
QY 362 ILSILLESVPELVCSMLHLCG-----TRLPAL 390  
DB 367 LLDIFLHEVNPSSLCGVIGLCARPELVEALEQAPATVSAALLKEPTPKQAPQKQAL 426  
QY 391 TVHTVQKDGFCCEVCKLVGYLDNLTENKSTKQIILALEKSGSFLLDPYQKODQFVA 450  
DB 427 PAHVPRQKNGFCCEVCKLVGYLDNLTENKSTKQIILALEKSGSFLLDPYQKODQFVA 486  
QY 451 EYEPVILTELVMDPSFVCLTIGACPSAHKPELLGTEKCIWGPSTWCONTEAQAQNAVE 510  
DB 487 EYEPVILTELVMDPSFVCLTIGACPSAHKPELLGTEKCIWGPSTWCONTEAQAQNAVE 546  
QY 511 HCKRHVWN 518  
DB 547 HCKRHVWN 554

RESULT 2  
Q7SY70 PRELIMINARY; PRT; 512 AA.  
AC Q7SY70.  
DT 01-OCT-2003 (TREMBLrel. 25, Created)  
DT 01-OCT-2003 (TREMBLrel. 25, Last sequence update)  
DT 01-OCT-2003 (TREMBLrel. 25, Last annotation update)  
DE Hypothetical protein.  
OS Xenopus laevis (African clawed frog).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidoidea; Pipidae;  
OC Xenopodidae; Xenopus.  
NC NCB1; TaxID=8355;  
LN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Whole;  
RX MEDLINE=22341132; PubMed=12454917;  
RA Klein S.U., Strausberg R.L., Wagner L., Pontius J., Clifton S.W.,  
RA Richardson P.,  
RT "genetic and genomic tools for Xenopus research: The NIH Xenopus  
RL Dev. Dyn. 225:384-391(2002).  
RN [2]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Whole;

RX MEDLINE=22388257; PubMed=12477932;  
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,  
RA Klausner R.D., Collins B.S., Wagner L., Shenmen C.M., Schuler G.D.,  
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,  
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Heien F.,  
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,  
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,  
RA Brownstein M.J., Ueda T.B., Tohiyuki S., Carninci P., Prange C.,  
RA Rana S.S., Lequellano N.A., Peters G.J., Abramson R.D., Mulhally S.J.,  
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,  
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,  
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,  
RA Fahy U., Heitlen E., Kettleman M., Madan A., Rodriguez S., Sanchez A.,  
RA Whitting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,  
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,  
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,  
RA Krzywinski M.I., Skalek U., Smalins D.E., Schmech A., Schein J.E.,  
RA Jones S.J., Marra M.A.,  
RT "Generation and initial analysis of more than 15,000 full-length human  
RT and mouse cDNA sequences.";  
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).  
RN [3]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Whole;  
RX Klein S., Strausberg R.,  
RL Submitted (JUL-2003) to the EMBL/GenBank/DBJ databases.  
DR EMBL; BC054988; AAH54988.1; -  
KW Hypothetical protein.  
SQ SEQUENCE 512 AA; 57555 MW; 57CFA5E2093F6FB1 CRC64;  
Query Match 52.1%; Score 1437; DB 13; Length 512;  
Best Local Similarity 46.6%; Pred. No. 6e-101;  
Matches 240; Conservative 130; Mismatches 131; Indels 14; Gaps 7;  
QY 5 LGALAGPVLGKCTRGSAVWCONVKTASDCGAVHGCLOTWNNKPVVSLPDDICKQV 64  
DB 11 LALVATPLFGRGCKAKPEVCEVNRTHSGAVHGCQOSVWNNKPVVSLPDDICKETI 70  
QY 65 TAAAGMKNATBEELIVLEKTCMDLPRKNMSASCKEIVDSYLPYILDIKGMERPE 124  
DB 71 TVLGNMKNITIGETIKYLVKVCVLPDPGILATCKQEVDSYLVNLLQDELSSNGV 130  
QY 125 VCSALNLCESLQKHAELNHOKEKSKIPELDMTEVAVPAPNITLILYPODGERSRQ 184  
DB 131 LCSSLGLCTSLQKHAELNHOKEKSKIPELDMTEVAVPAPNITLILYPODGERSRQ 188  
QY 185 PKQKXNDGVQDCIQWVTDIQTAVRTNSTFVQALVEHVEKEDCDRLGPGMADICKNYISQ 244  
DB 189 PK-TGDIQNDCTTLLSDVQDALRSNSFSKTLVDHFQECNLDPYAEKCKSYINQYAD 247  
QY 245 IAIOMMMHMOPEKICLVGFCDEVKEMQTLVPAKVASKNVLPALTEVPIKHEVPAK 304  
DB 248 IAIQVLMQKQKQCGVAGFCDEKSTPLQNIIPA---KSLIYAVVITE---NPLP-E 298  
QY 305 SDVYCEVCEFLVETKLIIDNNKTEKILIDADPKMSCKPKSLSECCQEVDTYSSILS 364  
DB 307 QAHNVILCQTCQFVNMKSEELIVNATBELVAKGLSNACALLPDPARTKCEVGTGFP 366  
QY 362 ILSILLESVPELVCSMLHLCG-----TRLPAL 390  
DB 367 LLDIFLHEVNPSSLCGVIGLCARPELVEALEQAPATVSAALLKEPTPKQAPQKQAL 426  
QY 391 TVHTVQKDGFCCEVCKLVGYLDNLTENKSTKQIILALEKSGSFLLDPYQKODQFVA 450  
DB 427 PAHVPRQKNGFCCEVCKLVGYLDNLTENKSTKQIILALEKSGSFLLDPYQKODQFVA 486  
QY 451 EYEPVILTELVMDPSFVCLTIGACPSAHKPELLGTEKCIWGPSTWCONTEAQAQNAVE 510  
DB 487 EYEPVILTELVMDPSFVCLTIGACPSAHKPELLGTEKCIWGPSTWCONTEAQAQNAVE 546  
QY 511 HCKRHVWN 518  
DB 547 HCKRHVWN 554

RESULT 3  
Q8UVZ4

ID	SEQUENCE	PRELIMINARY:	PRT:	520 AA.
AD	Q8UWZ4			
AC	Q8UWZ4			
DT	01-MAR-2002 (TREMBLrel. 20, Created)			
DT	01-MAR-2002 (TREMBLrel. 20, Last sequence update)			
DT	01-OCT-2003 (TREMBLrel. 25, Last annotation update)			
DB	Lysosomal cofactor/neurotrophic factor prosaposin.			
OS	Brachydanio rerio (zebrafish) (Danio rerio).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;			
OC	Cyprinidae; Danio.			
OX	NCBI_TaxID=7955;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RA	Seo H.-C., Lie O., Fjose A., O'Brien J.S., Kishimoto Y.;			
RL	"Cloning, expression and promoter analysis of zebrafish prosaposin.";			
RL	Submitted (JUN-2000) to the EMBL/Genbank/DBJ databases.			
DR	EMBL; AF267996; AAL54381.1; -			
DR	GO; GO:0005764; C:Lysosome; IEA.			
DR	GO; GO:0006665; P:sphingolipid metabolism; IEA.			
DR	InterPro; IPR003119; SAPA.			
DR	InterPro; IPR007856; SAPB_1.			
DR	InterPro; IPR008138; SAPB_2.			
DR	InterPro; IPR008140; SAPB_sub.			
DR	InterPro; IPR008373; SAPosin.			
DR	InterPro; IPR008139; SAPosinB.			
DR	Pfam; PF02199; SAPA; 2.			
DR	Pfam; PF05184; SAPB_1; 3.			
DR	Pfam; PF03489; SAPB_2; 4.			
DR	PRINTS; PRO1797; SAFOSIN.			
DR	ProDom; PD001732; SAPB_sub; 3.			
DR	SMART; SM00162; SAPA; 2.			
DR	SMART; SM00118; SAPB; 4.			
QO	SEQUENCE 520 AA; 57451 MW; F95620F84BA41CB5 CRC64;			

Query Match	Score	DB 13:	Length	520:
Best Local Similarity	47.0%	Pred. No. 4,9e-99		
Matches	246:	Conservative 120:	Mismatches 139:	Indels 18: Gaps 8
QY	4	LLGALAGPVLTGKECTRGSAVMCCNVNTASDCGAVKELQTVMNKPVTYSKPCDICKDV	63	
DB	8	LVTTAVASPLLTGTGCGARPPYWCQNVNTATSLCGAVHQCGQVNMNKKPQMKVTPCDLCEV	67	
QY	64	VTAAGDMKDNATBEELVYLEKTCIDMLPKPMMSACEIYDSYLPVILDIYKGMSPRG	123	
DB	68	LVVVEQLKDNVTESELGYLEKACQLPDEGLANQCKEIVDNPVPPVWGIIQGLIDBPG	127	
QY	124	EVCASALNCESLQKRLAELNMQKOLENSKIPBLDMTEVAPPMANIPILLYPQDPRSKP	183	
DB	128	VVCGALGCVSQQALAA---KAQLTSLNTEIPVDNLQRVSPLNIPOLLYEE---KRE	180	
QY	184	QPKDNDVCODCIQWATDIIQTAVRTNSFFVQALVHEVEECRLSPGMADICAKNTISOYS	243	
DB	181	TPKQKQDVCDQCVTTISPTQDEARVNSSFINTLINQVENEGLLPGMSDMCKEITISQYG	240	
QY	244	EIALQMMHMQPKELCALVFGC-DEVKEMPMQTLVPAK-VASKNVI PALELVEPI-----K	297	
DB	241	PLVFQQLMSMQPKDICABAGFCPTKQKSVPMKELPAKSIPAVKNFPAVKEKVPATMPA	300	
QY	298	KHEVPASADVCEVEEFLVKEVETKILDNKTKELILDAFDRKSCGLPKSLSECOEVDVT	357	
DB	301	KNLVAVRSPQALICEYVMKEIENNIQDOTSBAELVQAVEKCNILPSTLTQCKDLLET	360	
QY	358	YGSSTILSILLESVPELVCSMLHLCSG-TRLPAL-TVNHTPQKDSGFCVCKGLVGYLDR	415	
DB	361	YGOATIIDLVGEADPKTYCSFLATLCSGYSHVVMQKHFA---GQFCVCCMAVRYVDG	417	
QY	416	NLEKNSTKOEILIALEKSGSLPDEPYQKODQFVAEYEBVLLEILVENWDSFVCLKIGA	475	
DB	418	ILEQVATQSEIBEAVIKVCSFLPDVAWKECNCQILRQYEBLVAQLLQTLDPDFVCMKIGA	477	
QY	476	CPSAKHPLGTREKCIWGPSYWCQNTETLAQCAVAEHCKRHVNN	518	
DB	478	CPEAVORLLGNQCSWGPAYMCKNVTQARCNALHCRCHVNS	520	

SEQUENCE	522 AA;	57671 MW;	D3C15A305725C1CD CRC64;
09DGB2	PRELIMINARY;	PRT;	522 AA.
AC	09DGB2;		
DT	01-MAR-2001	(TREMBLrel. 16, Created)	
DT	01-MAR-2001	(TREMBLrel. 16, Last sequence update)	
DT	01-OCT-2003	(TREMBLrel. 25, Last annotation update)	
DE	Prosaposin.		
GN	PSAP.		
OS	Birchydanio rerio (zebrafish) (Danio rerio).		
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
OC	Actinopterygii; Neopterygii; Teleostei; Osteichthyes; Cypriniformes;		
OC	Cyprinidae; Danio.		
OX	NCBI_TaxID=7955;		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RA	Altman N., Horowitz M.;		
RT	"The zebrafish prosaposin cDNA."		
RL	Submitted (NOV-1998) to the EMBL/GenBank/DBJ databases.		
DR	EMBL; AF108655; AAC32919.1; -.		
DR	ZFIN; ZDB-GENE-020108-1; psap.		
DR	GO; GO:0005764; C:lysosome; IEA.		
DR	GO; GO:0006665; P:sphingolipid metabolism; IEA.		
DR	InterPro; IPR003119; SApA.		
DR	InterPro; IPR007856; SApB_1.		
DR	InterPro; IPR008138; SApB_2.		
DR	InterPro; IPR008140; SApB_sub.		
DR	InterPro; IPR008373; Saposin.		
DR	InterPro; IPR008139; SaposinB.		
DR	pfam; PF02199; SApA_2.		
DR	pfam; PF05184; SApB_1; 3.		
DR	pfam; PF03489; SApB_2; 4.		
DR	PRINTS; PR01797; SApOSIN.		
DR	Prodom; PD001732; SApB_sub; 3.		
DR	SMART; SM00152; SApA_2.		
DR	SMART; SM00118; SApB; 4.		
SO	SEQUENCE	522 AA;	57671 MW;
			D3C15A305725C1CD CRC64;

Query	March	Similarity	48.6%	Score 1341	DB 13	Length 522
Best Local	Similarity	45.7%	Pred. No. 1.2e-93			
Matches 240	Conservative 116	Mismatches 149	Indels 20	Gaps 9		
QY	4	LIGALAGPVLGLECTRGSAMWCQNVKTA	DCGAVRHCLQTWNKPKTVKSLPCDICKOV	63		
DB	8	LVTVAVASPILTGTBQCARGPYWCQNVKTA	SLGAVOHCCQNVNKKPKMKTVPICDLCKEV	67		
QY	64	VTAAGDMVLKDNATSEELVLTVEKTCMDL	PKPNNMSASCKEIVDSTLPIVLTIDIKEMSRPG	123		
DB	68	LWVEQLKKNVTESELIGYLEKACCOLI	PBEGLANCKEIVTTSISOPSMASCKELDDPG	127		
QY	124	EVCSSALNLCESLQGLAEINLHOKOLESNK	TELPDMTEVAFPMANIPILTVPODGPRSKP	183		
DB	128	VVCALGELGCSQQAALA----	KQDLSNEIPQVDLNGRVSPFLNITPOLLYPE--	180		
QY	184	QPKDNGVCCDDCIOMVTDIQTAVTNS	TFVQALVEHYKEECDRLGPGMADICKNYISOYS	243		
DB	181	TPKQGHVCCDCCVFIFDIODEAVN	SNSSFITLLTIAOVENOCCELLGPGMSDMCKEYISYG	240		
QY	244	EIALQMMMHM--OPKEICALVGFC-	DEKEMPMOTVLPAK-VASKNYIPALLETVEPI---	296		
DB	241	PLVPOQLMSHQQKDCIARAGFCPT	QKQSVPMKPLPAKSI PAVKMFPAVKEKVPVATM	300		
QY	297	KKHEVPAKSDVYCEVCFLWKEVYTKLI	DNNKTEKEIILDAFDKMSCKLPLKSLSECCQEV	355		
DB	301	PAKMLVVRSDSPQCAICEYMKKEIEN	NIIOOTSSEAIYQAVEKCNILIPSTLTQCKDLI	360		
QY	356	DTYSSSLTILTEVSEPEIVCSMHLTSG-	TRLPAL-TVHVTQPKDGFCEVCCKLGYL	413		
DB	361	ETVQDAILTDLVQADPPTVCSFLATCS	GVSHVEMDKQHPAA--GGFCVCCMAARVY	417		
QY	414	DRNLKSTKQKELLALAEKSGSLPDP	QOKCDDFVAEYEVLTLEIIVENVDPSPVCLKI	473		

Db 418 DGLFQNTQSGIEEAVLKVCSFLPYAKDSECONLTIEEYELPLVQLLTQDLPDVCMKL 477  
Qy 474 GACGPAHAPLIGTEKCIWPSYWCNTETAAQCAVAHECKRHWN 518  
Db 478 GACPEAVQRLIGLNCQSWGPAWCKNVGTARCNALNCRHWN 522

RESULT 5  
OBC1C1 PRELIMINARY; PRT; 525 AA.  
ID OBC1C1  
AC OBC1C1  
DT 01-MAR-2003 (TRENBLrel. 23, Created)  
DT 01-MAR-2003 (TRENBLrel. 23, Last sequence update)  
DT 01-OCT-2003 (TRENBLrel. 25, Last annotation update)  
DE Hypothetical saposin A-type domain/saposin type B containing protein.  
GN 2310020A21RIK.  
OS Mus musculus (Mouse).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
OX NCBI\_TaxID=10090;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=C57BL/6J; TISSUE=Skin;  
RX MEDLINE=22354683; PubMed=12466851;  
RA The FANTOM Consortium,  
the RIKEN Genome Exploration Research Group Phase I & II Team;  
RT "Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs."  
RL Nature 420:563-573(2002).  
DR EMBL; AK028455; BAC25861.1; -  
DR MGD; MGI:1924193; 2310020A21RIK.  
DR GO; GO:0005764; C:lysosome; IEA.  
DR GO; GO:0006665; P:sphingolipid metabolism; IEA.  
DR InterPro; IPR003119; Sapa.  
DR InterPro; IPR007856; Sapa\_1.  
DR InterPro; IPR008138; Sapa\_2.  
DR InterPro; IPR008140; Sapa\_sub.  
DR InterPro; IPR008373; Saposin.  
DR InterPro; IPR008139; SaposinB.  
DR Pfam; PF02199; SAPA; 2.  
DR Pfam; PF05184; Sapa\_1; 4.  
DR Pfam; PF03489; Sapa\_2; 4.  
DR PRINTS; PR01797; Saposin.  
DR ProDom; PD001732; Sapa\_sub; 3.  
DR SMART; SM00162; SAPA\_2.  
DR SMART; SM00118; Sapa; 4.  
KW Hypothetical protein.  
SQ SEQUENCE 525 AA; 57350 MW; 374F6050CD4D223 CRC64;

Query Match 39.5%; Score 1089.5; DB 11; Length 525;  
Best Local Similarity 40.0%; Pred. No. 1.6e-74;  
Matches 211; Conservative 108; Mismatches 172; Indels 37; Gaps 9;

Qy 2 ASLIGALAGVYLGLKECTRSAAVWCQNVKTAQDCGAVKCLQTWNKPTVYSLPCDICK 61  
Db 12 SGLIGAAASPIISVRECAKSEVWCQDLQAAAKCAVHRCQSAVWKKPTVYSLPCSCVQ 71  
Qy 62 DVVTAAGDMKDNATTEELIVLEKTCMDLPKPNMSASKEIYDSYLPVLLDIKGEKMR 121  
Db 72 DVAAAAGNGVNPAGTIESDILTIVMKTCEWLPSQESSAKCKMNVNNSAVALSMLSGAQT 131  
Qy 122 P-GEVCSALNLCESLQGHIAELNKHQKQLESNKIPELDMTEVVA PPMANIPLLIYPDGR 180  
Db 132 DLASVCTALTLCEPLQGHIAELTTSERLPLOE-----DANEVMA PPLISNGLSFHPQMP- 185  
Qy 181 SKPQKNGVYCODCIQMTDIDQAVFNTSFVQALVEHYKECDRLGGMADICKNYIS 240  
Db 186 -----EGAVCHDCVQVLIISLDALLESNTLTLAEVTQN---QCGMGGLALCENTYH 235  
Qy 241 QVSEIALQMMHMQPKICALVGFCEVEMPMQTLVPAKVASKVNTPALLEVEPIKQHE 300

Db 236 RQFVPAKQTLQGLPPQEVCKRGCEERSAHWL-----TRVAADVPSLSEMEWP-RTNE 289  
Qy 301 VPAKSDVYCEVCEFLVKEVNTLIDNNKTEKEILDAPDKMSKLPKSLSECEQVDPYGS 360  
Db 230 LQMOGLTCDCVCLNLDVDELKMLVTNSTEALISHTLERVCTVPEPLVQCITLVDTISP 349  
Qy 361 SILSILIEEVSPELVCSMLHCSGTR-----LPLVTVHTVTPKQDGFCEVCKKL 409  
Db 350 ELVQ-LMSKVTPEVVCETIKLCSKRRARISRAVATPPLP--VDEENQSFQCCCKRL 406  
Qy 410 VGYIDRLNKNSTKQEIILAEKGSFLPDPYQKCCQDFVAYEVPVILITLVEVMPSPV 469  
Db 407 LQMSQNDLHKSTRTDILNFKGCCRILPLPYWQCNRFPAEYEPVILISLKFMMNPDL 466  
Qy 470 CLKIGAPSAHAPLIGTEKCIWPSYWCNTETAAQCAVAHECKRHWN 517  
Db 467 CKMGAGCHGKRTPLGTQCVWGPSFWCKSPFAALECNALHQCRLW 514

RESULT 6  
OBBVJS PRELIMINARY; PRT; 449 AA.  
ID OBBVJS  
AC OBBVJS  
DT 01-MAR-2003 (TRENBLrel. 23, Created)  
DT 01-MAR-2003 (TRENBLrel. 23, Last sequence update)  
DT 01-OCT-2003 (TRENBLrel. 25, Last annotation update)  
DE Hypothetical saposin A-type domain/saposin type B containing protein.  
GN 2310020A21RIK.  
OS Mus musculus (Mouse).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
OX NCBI\_TaxID=10090;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=C57BL/6J; TISSUE=Eye;  
RX MEDLINE=22354683; PubMed=12466851;  
RA The FANTOM Consortium,  
the RIKEN Genome Exploration Research Group Phase I & II Team;  
RT "Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs."  
RL Nature 420:563-573(2002).  
DR EMBL; AK078699; BAC37363.1; -  
DR MGD; MGI:1924193; 2310020A21RIK.  
DR GO; GO:0005764; C:lysosome; IEA.  
DR GO; GO:0006665; P:sphingolipid metabolism; IEA.  
DR InterPro; IPR003119; Sapa.  
DR InterPro; IPR007856; Sapa\_1.  
DR InterPro; IPR008138; Sapa\_2.  
DR InterPro; IPR008140; Sapa\_sub.  
DR InterPro; IPR008373; Saposin.  
DR InterPro; IPR008139; SaposinB.  
DR Pfam; PF02199; SAPA; 1.  
DR Pfam; PF05184; Sapa\_1; 3.  
DR Pfam; PF03489; Sapa\_2; 3.  
DR PRINTS; PR01797; Saposin.  
DR ProDom; PD001732; Sapa\_sub; 2.  
DR SMART; SM00162; SAPA\_1.  
DR SMART; SM00118; Sapa; 3.  
KW Hypothetical protein.  
SQ SEQUENCE 449 AA; 48705 MW; 377CC42A475B292F CRC64;

Query Match 28.1%; Score 776; DB 11; Length 449;  
Best Local Similarity 35.9%; Pred. No. 9.4e-51;  
Matches 165; Conservative 90; Mismatches 160; Indels 44; Gaps 12;

Qy 2 ASLIGALAGVYLGLKECTRSAAVWCQNVKTAQDCGAVKCLQTWNKPTVYSLPCDICK 61  
Db 12 SGLIGAAASPIISVRECAKSEVWCQDLQAAAKCAVHRCQSAVWKKPTVYSLPCSCVQ 71  
Qy 62 DVVTAAGDMKDNATTEELIVLEKTCMDLPKPNMSASKEIYDSYLPVLLDIKGEKMR 121  
Db 72 DVAAAAGNGVNPAGTIESDILTIVMKTCEWLPSQESSAKCKMNVNNSAVALSMLSGAQT 131

QY 122 P-GEVCSALNLCESLQKHLAEINHOKELESNKIPELDMTEVAVPMPANIPLLLYPODGR 180  
DB 132 DLASCTALTTLCEPIQRHAEITTSERPLTQE-----DANEVAPFLSNALSFHSGMP- 185  
QY 181 SKPOKNDGVQDCDIQWVTDIQTAVRTNSTFVQALVEHVKEECRLGPMADICKNYIS 240  
DB 186 -----BGAVCHDCVQIISLLQDALLESNTLAEVTVQN---QCSMGGLAALCENYIH 235  
QY 241 QYSEIATQMOMQKEICATLVGFCDEVKEMMOTLVPAKVASKVITPALVEPIKHEVPAK 300  
DB 236 RQFVPAKQTLQGLPQOEVCCKGFCERESAHWL-----TRVAAVGVPSLEMEMB-RTNE 289  
QY 301 VPAKSDVYCEVCEFLVKEVETKLIIDNNKTEKEILDAPDKMCSLPSKISECOEVDYTS 360  
DB 290 LQWQAGQTCVCLNVQELDKMLVYNSTALISHLERCTVAPPELVQCCITLVDTISPELVQ 349  
QY 361 SILSEVSPELVCSMLHLCSGTR-----LPALTVHTYTOPKDGFCCEVCKL 409  
DB 350 ELVQ-LMSKVTEPEKVCETIKLCSRRARSISRVAVATPSP--VDEBNQSGFCQCKRL 406  
QY 410 VGYIDRNLKSTKQELIALLKGC--SFLPPYQKCD 446  
DB 407 LGH----VFPSGPOEHQAG--HSECLQRLPDPVALCD 440

## RESULT 7

Q8CIN0 PRELIMINARY; PRT; 402 AA.  
AC Q8CIN0;  
DT 01-MAR-2003 (TREMBLrel. 23, Created)  
DT 01-MAR-2003 (TREMBLrel. 23, Last sequence update)  
DT 01-OCT-2003 (TREMBLrel. 25, Last annotation update)  
DE Hypothetical saposin A-type domain/saposin type B containing proteain.  
GN 2310020A2IRIK.  
OS Mus musculus (Mouse).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
OX NCBI\_TaxID=10090;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=C57BL/6J; TISSUE=Tongue;  
RX MEDLINE=22354683; PubMed=12466851;  
RA The RIKEN Genome Exploration Research Group Phase I & II Team;  
RA "Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs."  
RT Nature 420:563-573(2002).  
RL EMBL; AK009408; BAC25258.1;  
DR MGD; MGI:1924193; 2310020A2IRIK.  
DR GO; GO:0005764; C:lysosome; IEA.  
DR GO; GO:0006665; P:sphingolipid metabolism; IEA.  
DR InterPro; IPR003119; SAPA.  
DR InterPro; IPR007856; SAPB.1.  
DR InterPro; IPR008138; SAPB.2.  
DR InterPro; IPR008140; SAPB\_sub.  
DR InterPro; IPR008373; SaposinB.  
DR InterPro; IPR008139; SaposinB.  
DR Pfam; PF0184; SAPB.1.  
DR Pfam; PF0184; SAPB.1.  
DR Pfam; PF03489; SAPB.2.  
DR PRINTS; PR01797; SAFOSIN.  
DR ProDom; PD001732; SAPB\_sub.2.  
DR SMART; SM00162; SAPA.1.  
DR SMART; SM00118; SAPB.3.  
KW Hypothetical protein.  
SQ SEQUENCE 402 AA; 44420 MW; E90017CBF4017ED6 CRC64;

Query Match 28.0%; Score 772; DB 11; Length 402;  
Best Local Similarity 38.1%; Pred. No. 1.6e-50;  
Matches 154; Conservative 85; Mismatches 129; Indels 36; Gaps 8;

QY 125 VCSALNLCESLQKHLAEINHOKELESNKIPELDMTEVAVPMPANIPLLLYPODGRSKRQ 184  
DB 13 VCTALTTLCEPIQRHAEITTSERPLTQE-----DANEVAPFLSNALSFHSGMP- 62  
QY 185 PRKNDGVQDCDIQWVTDIQTAVRTNSTFVQALVEHVKEECRLGPMADICKNYISQYSE 244  
DB 63 ---BGAVCHDCVQIISLLQDALLESNTLAEVTVQN---QCSMGGLAALCENYIHQFV 116  
QY 245 IALQMMQMOKEICATLVGFCDEVKEMMOTLVPAKVASKVITPALVEPIKHEVPAK 304  
DB 117 PAKQTLQGLPQOEVCCKGFCERESAHWL-----TRVAAVGVPSLEMEMB-RTNE 170  
QY 305 SDVYCEVCEFLVKEVETKLIIDNNKTEKEILDAPDKMCSLPSKISECOEVDYTS 364  
DB 171 LGITDVCCLNVQELDKMLVYNSTALISHLERCTVAPPELVQCCITLVDTISPELVQ 230  
QY 365 ILSEVSPELVCSMLHLCSGTR-----LPALTVHTYTOPKDGFCCEVCKL 413  
DB 231 -LMSKVTEPEKVCETIKLCSRRARSISRVAVATPSP--VDEBNQSGFCQCKRL 287  
QY 414 DRNLKSTKQELIALLKGC--SFLPPYQKCD 446  
DB 288 SQNLDRKSTKQELIALLKGC--SFLPPYQKCD 446  
QY 474 GACPSAHKPLTGTEKCIWGPSYWCNTETAQCNVAHEKCRHW 517  
DB 348 GACHQPTPLTGTCVWGPSFWCKSPBALEMCNLAHCORLW 391

## RESULT 8

Q9Y125 PRELIMINARY; PRT; 953 AA.  
AC Q9Y125;  
DT 01-NOV-1999 (TREMBLrel. 12, Created)  
DT 01-NOV-1999 (TREMBLrel. 12, Last sequence update)  
DT 01-OCT-2003 (TREMBLrel. 25, Last annotation update)  
DE SAP-R protein.  
GN SAP-R OR BCDA:GH08312 OR CG12070.  
OS Drosophila melanogaster (Fruit Fly).  
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.  
OX NCBI\_TaxID=7227;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=Berkely;  
RX MEDLINE=20196006; PubMed=10731132;  
RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D., Amanatides P.G., Scherer S.E., Li P.W., Hoekstra R.A., Gallego J.P., George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N., Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X., Brandon R.C., Rogers Y.-H.C., Blazey R.G., Chapple M., Pfeiffer B.D., Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G., Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D., Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M., Beeson K.Y., Benos P.V., Bernier B.P., Bhandari D., Bolshakov S., Borkova D., Botchan M.R., Bouck J., Brokstein P., Brotlier P., Butts K.C., Buesam D.A., Butler H., Cadieu E., Center A., Chandra I., Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P., de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz J., Dodson K., Dou P.L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P., Durbin K.J., Evangelista C.C., Ferraz C., Ferrieres S., Fleischmann W., Foeller C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glaeser K., Glodek A., Gong P., Gorrell J.H., Gu Z., Guan P., Harris M., Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J., Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwan C., Jajani M., Kalush F., Karpen G.H., Ke Z., Kemisun J.A., Ketchum K.A., Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z., Laeko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X., Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D., Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A., Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L., Nelson D.R., Nelson K.A., Nixon K., Nuske D.R., Paclob J.M.,

RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,  
 RA Rehner K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,  
 RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,  
 RA Spler E., Spredling A.C., Stapleton M., Strong R., Sun E.,  
 RA Switzkes R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,  
 RA Wang Z.-Y., Wassarman D.A., Weinstein G.M., Weisenbach J.,  
 RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,  
 RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,  
 RA Zheng X.H., Zhong F.N., Zhou W., Zhou X., Zhu S., Zhu X., Smith H.O.,  
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.,  
 RT "The genome sequence of *Drosophila melanogaster*";  
 RL Science 287:2185-2195(2000).  
 RN  
 RP SEQUENCE FROM N.A.  
 RA Rubin G.M., Wan K.H., Harvey D., Lewis S.E., Broksrein P., Tsang G.,  
 RA Agbayani A., Arcaina T.T., Baxter E., Blazey R.G., Butenloft C.,  
 RA Champe M., Chavez C., Chew M., Doyle C.M., Farfan D.E., Frise E.,  
 RA Galle R., George R.A., Harris N.L., Hoskins R.A., Evans-Holm M.,  
 RA Houston K.A., Hummel S.R., Kim E., Li P., Moshrefi M., Pacleb J.M.,  
 RA Park S., Seguelita A., Sethi H., Smit E., Switzkes R.R., Weinburg T.,  
 RA Ceiniker S.E.,  
 RT "Full length *Drosophila melanogaster* cDNA sequence";  
 RL Submitted (Apr-1999) to the EMBL/GenBank/DBJ databases.  
 DR EMBL: AEO03775; AAF57097.1; -  
 DR EMBL: AF145647; AAD38622.1; -  
 DR FlyBase: FBgn0000416; Sap-t.  
 DR GO: GO:0005764; C:lysosome; IEA.  
 DR GO: GO:0006665; P:sphingolipid metabolism; IEA.  
 DR InterPro: IPR003119; Sapa.  
 DR InterPro: IPR007856; Sapa.1.  
 DR InterPro: IPR008138; Sapa.2.  
 DR InterPro: IPR008140; Sapa\_sub.  
 DR InterPro: IPR008373; Saposin.  
 DR InterPro: IPR008139; Saposinb.  
 DR Pfam: PF02189; Sapa.1.  
 DR Pfam: PF05184; Sapa.1; 6.  
 DR Pfam: PF03489; Sapa.2; 7.  
 DR PRINTS: PRO1797; SAPOSIN.  
 DR ProDom: PD001732; Sapa\_sub; 6.  
 DR SMART: SM00162; SAPA\_1.  
 DR SMART: SM00118; SAPA; 7.  
 SO SEQUENCE 953 AA; 105962 MW; D6CFD3E9D1502A8 CRC64;  
 Query Match 19.9%; Score 549.5; DB 5; Length 953;  
 Best Local Similarity 23.2%; Pred. No. 4e-33;  
 Matches 144; Conservative 105; Mismatches 212; Indels 161; Gaps 18;  
 QY 8 ALAGVGLIKECTRGSAVWCQNVKTASDCGAVKICLTQVWNR---PTVSKPDCICKDV 64  
 DB 20 AATATLASSKCTWGPSTWCGNFSNKSKECRATRHCITQVWETQKVPVDTISICTCKDV 79  
 QY 65 TAAGDMLKDNATTEETIIVLYLEKTCQMLPKPMMSASCKEIVDSYIPVLDTIKGMSRPG 124  
 DB 80 TQARQQLSNQTEETELKEVFEGSKCLIPKIQKECTIKVADDFELVEALASQMN-PDQ 138  
 QY 125 VCSALNLCES-----LQKH-----LAELN 143  
 DB 139 VCSVAGLCNSARIDELYNKGIQAGLDGVQNEDESSSETELAMQPNLSGNCNLLSRIM 198  
 QY 144 HOKQLESKIPDELMTETV-----APFMANIPLLI----- 173  
 DB 199 HSKRAADRR--DMVETMLHMGCSLSSFSDACANITLVTFENDIYDHSKELTTDAVCHV 255  
 QY 174 -----YPDGPRSKPOPK--DNQD-----VCQDCIOMVTDIQTAVTNSITFQALV 217  
 DB 256 SGVCASHYHOHEEKPOEALVLMADAGDIPCELCQGLVGLADVAVANTTEFEFQVWE 315  
 QY 218 EHYVEECDRIGPGMADICKNYISQYSRTAIQ-MMMHMQPREICALVGFDE-----VKEM 271  
 DB 316 GFCKQ-----SKGFCKECLISIVQYHVIETLVSKLDNAGACMIGICQKXASASMKOV 370  
 QY 272 PMQTLVPAKVASKIVPALEL---VEPIKKE-----VP----- 302

DB 371 PIMPLP-----VIEPAQKITTEKLEKEKQKOLASPEKRSQOEITLDMQPIHLMG 423  
 QY 303 -----AKSDVYCEVECFVKEVTKLIDNNKEKEIILDAFDKMCSPKPSISECOEV 355  
 DB 424 NAMPALVEGELTTLCEYMHFIQETLATPSTDEIKHVENICATLPSGVAGOCRNFY 483  
 QY 356 DYSSISILIEEVSPELVCSMLHCSGTRLPALTYHTQP-----KDGFEVCKKL 409  
 DB 484 EMTGAVATIALVQGINRDVCPVLMQCPKPLPKXEDVEVNPQASDEOPPTCPCLPA 543  
 QY 410 VGYIDRLNEKXSTKOEIILAEKCSFLPDPYQKQCFVAEVEPVILIEIVEMPPSV 469  
 DB 544 VEOQMIRNDKSKDKNTKVTNGLCSHLPEIKEECVDPFNTYSNLLIMLITDFRQEI 603  
 QY 470 CLKTAGCSAMKPLGKEKCIW 491  
 DB 604 CVQKLCPR-----TTVALW 618

## RESULT 9

Q8N7T4

ID Q8N7T4

AC Q8N7T4

DT 01-OCT-2002 (TrEMBLrel. 22, Created)

DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)

DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)

DE Hypothetical protein FLJ40379.

OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.

OX NCBI\_TaxId=9606;

RN [1]

RP SEQUENCE FROM N.A.

RC TISSUE=Testis;

RA Oshima A., Takahashi-Fujii A., Tanase T., Imose N., Takeuchi K.,

RA Arita M., Musashino K., Yunki H., Hara H., Sugiyama T., Irie R.,

RA Otsuki T., Sato H., Wakamatsu A., Ishii S., Yamamoto J., Isono Y.,

RA Kawai-Hio Y., Saito K., Nishikawa T., Kimura K., Yamashita H.,

RA Matsuo K., Nakamura Y., Sekine M., Kikuchi H., Kanda K., Magatsuma M.,

RA Murakawa K., Kanehori K., Sugiyama A., Kawakami B., Suzuki Y.,

RA Sugano S., Nagahori K., Maehiro Y., Nagai K., Tsogai T.,

RT "NEO human cDNA sequencing project";

RL Submitted (JUL-2002) to the EMBL/GenBank/DBJ databases.

DR EMBL: AK097698; BAC05143.1; -

DR GO: GO:0005764; C:lysosome; IEA.

DR GO: GO:0006665; P:sphingolipid metabolism; IEA.

DR InterPro: IPR003119; Sapa.

DR InterPro: IPR007856; Sapa.1.

DR InterPro: IPR008138; Sapa.2.

DR InterPro: IPR008140; Sapa\_sub.

DR InterPro: IPR008373; Saposin.

DR InterPro: IPR008139; Saposinb.

DR Pfam: PF02189; SAPA.1.

DR Pfam: PF05184; Sapa.1; 2.

DR Pfam: PF03489; Sapa.2; 2.

DR PRINTS: PRO1797; SAPOSIN.

DR ProDom: PD001732; Sapa\_sub; 2.

DR SMART: SM00162; SAPA\_1.

DR SMART: SM00118; SAPA; 2.

KW Hypothetical protein.

SO SEQUENCE 241 AA; 26719 MW; BE29EAF6EE6E6AB CRC64;

Query Match 19.8%; Score 545; DB 4; Length 241;  
 Best Local Similarity 44.6%; Pred. No. 1.6e-33;  
 Matches 107; Conservative 45; Mismatches 76; Indels 12; Gaps 4;

QY 287 IPALVEPIKKEHVAKSVYCEVFLYKVTYKLIIDNNKEKEIILDAFDKMCSPKPS 346  
 DB 4 VPSDELGLFPRKQSEWQKGVTCVCMNVQKLDHLMKNSSELMITHALERVCVMPAS 63  
 QY 347 LSEECQVWVTYSSSILIEEVSPELVCSMLHCSGTRLPALTYHTQ-----VTPK--- 398  
 DB 64 ITKECIIIVDTYSPSLVQ-LVAKITPEKVCKFIKRLC-GNRRRAVRAVDAVATVPSPEWDA 121

QY 399 --DGGFCVCKLVGYLDRLNLEKSTKOEIIAALFKGCSFLPDPYKODQDQVFAVEPEVL 456  
DB 122 ENQSGFCNGCKRLTVSSSHNLSKSTKRIILVAFKGGCILEPVMIOCKHVTQYIEPVL 181  
QY 457 IEILVEVNDPSFVCLIKIGACPSAHRPLLTGTEKCIKNGPSYWCNTETAACNAVECHCKRY 516  
DB 182 IESLMDMNDPVAVCCKKVGACHPRTPLIGTDCALGSPFWCSQGAALKCNVHCQKHV 241

RESULT 10  
ID P79254 PRELIMINARY; PRT; 121 AA.  
AC P79254;  
DT 01-MAY-1997 (TREMBlrel. 03, Created)  
DT 01-MAY-1997 (TREMBlrel. 03, Last sequence update)  
DT 01-OCT-2003 (TREMBlrel. 25, Last annotation update)  
DE Sulfated glycoprotein-1/SGP-1 (Fragment).  
OS Ovis aries (Sheep).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;  
OC Bovidae; Caprinae; Ovis.  
OX NCBI\_TaxID=9940;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=96249303; PubMed=8648570;  
RA Spencer T.E., Graf G.H., Bazer F.W.;  
RT "Sulfated glycoprotein-1 (SGP-1) expression in ovine endometrium  
during the oestrous cycle and early pregnancy."  
RL Reprod. Fertil. Dev. 7:1053-1060(1995).  
DR EMBL; S82555; AAD14405.1; -.  
DR GO; GO:0005764; C:lysosome; IEA.  
DR GO; GO:0006665; P:sphingolipid metabolism; IEA.  
DR InterPro; IPR008138; SAPB\_2.  
DR InterPro; IPR008373; Saposin.  
DR InterPro; IPR008139; SaposinB.  
DR Pfam; PF03489; SAPB\_2; 1.  
DR PRINTS; PR01797; SAFOSIN.  
DR SMART; SM00118; SAPB; 1.  
FT NON\_TER 1  
SQ SEQUENCE 121 AA; 13604 MW; 4F0F5A6EB83D0C9A CRC64;

Query Match 18.7%; Score 517; DB 6; Length 121;  
Best Local Similarity 77.7%; Pred. No. 9e-32; Indels 0; Gaps 0;  
Matches 94; Conservative 18; Mismatches 9; Indels 0; Gaps 0;

QY 199 VTDIOTAVRINSTFYQALVEHVEKCDRLGPGMADICKNYISQYSEIALQMMHMQPKXI 258  
DB 1 VTIDIOTAVRINSTFYQALVEHVEKCDRLGPGMADICKNYISQYSEIALQMMHMQPKXI 60

QY 259 CALVGFCDVEKEMQTLVPAKVASKNVILPALEVEPIKQHEVPAKSDVYCEVCEFLVKE 318  
DB 61 SALVGFCDVEKESPMRTLDPAKASKNILLPALETEPNKQKRI PAOTLIFCQVCQFVMBE 120

QY 319 V 319  
DB 121 V 121

RESULT 11  
ID O15997 PRELIMINARY; PRT; 965 AA.  
AC O15997;  
DT 01-JAN-1998 (TREMBlrel. 05, Created)  
DT 01-JAN-1998 (TREMBlrel. 05, Last sequence update)  
DT 01-OCT-2003 (TREMBlrel. 25, Last annotation update)  
DE Bmp109.  
OS Bombyx mori (Silk moth).  
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;  
OC Neoptera; Endopterygota; Lepidoptera; Glossata; Ditrysia; Bombycoidea;  
OC Bombycidae; Bombyx.  
OX NCBI\_TaxID=7091;  
RN [1]

RP SEQUENCE FROM N.A.  
RX MEDLINE=96279844; PubMed=9611271;  
RA Tambunan J., Chang P.-K., Li H., Natori M.;  
RT "Molecular cloning of a cDNA encoding a silkworm protein which  
contains the conserved BH regions of Bcl-2 family proteins."  
RL Gene 212:287-293(1998).  
DR EMBL; AB008449; BAA23126.1; -.  
DR PIR; T00207; T00207.  
DR GO; GO:0005764; C:lysosome; IEA.  
DR GO; GO:0006665; P:sphingolipid metabolism; IEA.  
DR InterPro; IPR003119; SAPA.  
DR InterPro; IPR007856; SAPB\_1.  
DR InterPro; IPR008138; SAPB\_2.  
DR InterPro; IPR008140; SAPB\_sub.  
DR InterPro; IPR008373; Saposin.  
DR InterPro; IPR008139; SaposinB.  
DR Pfam; PF02199; SAPA; 2.  
DR Pfam; PF05184; SAPB\_1; 6.  
DR Pfam; PF03489; SAPB\_2; 7.  
DR PRINTS; PR01797; SAFOSIN.  
DR ProDom; PD001732; SAPB\_sub; 6.  
DR SMART; SM00162; SAPA; 2.  
DR SMART; SM00118; SAPB; 7.  
SQ SEQUENCE 965 AA; 108825 MW; FA1A7BEETF626078 CRC64;

Query Match 18.5%; Score 511; DB 5; Length 965;  
Best Local Similarity 24.6%; Pred. No. 3.5e-30;  
Matches 139; Conservative 104; Mismatches 229; Indels 94; Gaps 18;

QY 13 VLGLKCTCRGSAVWCQNYKATSDCGAVGHCQTWNK---PTVASLPDICKDVYTAAGD 69  
DB 170 LLGKRCITWGPSPGNSNSTGECNATPHCINRWYSKTFPBDNDNIOICLIDVYKQARD 229

QY 70 MKDNATSEELVYLEKTCMLPKPNMSASCKEIVDSYLPVILDIKEMSPGEGYCAL 129  
DB 230 QLOSNETODEIKEYEVGSCKLIPKFVAGCMKLADFEVBLIETLASEMN-PQAVCSVA 288

QY 130 NLCS--LQKILAEINHOLES-----NKIPELDMTEVVAFFA----- 167  
DB 289 GLCNNAKIDRLLVYDVAQRELRAGCYNCQKTGVVRRKFDKTXEDFLVGLLYQVRNDS 348

QY 168 ---NIPLLVP-----ODG-----PRSKPQKNGD 190  
DB 349 LSDGSMILIFYKYLEILAVKDLNPSGICHVSGQSYKFNHDEFTPEQWQVSARD 408

QY 191 V-CODCIQMTVDIOTAVRINSTFYQALVEHVEKCDRLGPGMADICKNYISQYSEIALQ 249  
DB 409 VPCEFCEDLVNGLRVLANATTELE-FYKVLQGLCKQTKG-FKDECLHLAEQYVVIYNF 466

QY 250 MM-HMQPEICALVGFCDVEKEMQTLV-----PAKVA-----SKNVIP 288  
DB 467 LVSDLKPAETCKMIGICNLTSAPISPVARELVKVPKILGAESKTIARVPLAKOMEP 526

QY 289 ALLEVE--PIKQHEVPA-KSDVYCEVCEFLVKEVYKLLDNKTEKEIILDAPDKMSKLPK 345  
DB 527 ASAAVSVPLERMFLAQAQSKAACAFQYFLHYLOVLSDRTEKVKVAAQVQACDALPD 586

QY 346 SLSECEQEVVDYTGSSILSLLEEVSPELVCSMHLCSGTLPALTYHTVTPKDGFCFV 405  
DB 587 ALNGECKEFVQYGSAAVIALVQELDPASVCPALDICTGTS-EIRRVVNSEKSN--CPL 643

QY 406 CKLVGYIDRLNLEKSTKOEIIAALFKGCSFLPDPYKODQDQVFAVEPEVLIEILVEVND 465  
DB 644 CLFAVEQLESVLKNNRSEENIRKALDGLCTRLSQDQSECIIDFVYSSQLVEMLVADMN 703

QY 466 PSFVCLIKIGAC-PSAHRPLLTGTEKCI 490  
DB 704 AKETCVPLKCRDQLHDLPLKLTHTSSI 729

RESULT 12  
ID O8IMH4 PRELIMINARY; PRT; 876 AA.  
AC O8IMH4  
DT 08IMH4



AC Q81MH4;  
 DT 01-MAR-2003 (Tremblrel. 23. Created)  
 DT 01-JUN-2003 (Tremblrel. 24. Last sequence update)  
 DT 01-OCT-2003 (Tremblrel. 25. Last annotation update)  
 DE CG12070-PB.  
 CN SARP-R.  
 OS Drosophila melanogaster (fruit fly).  
 OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;  
 OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;  
 OC Ephydriidae; Drosophilidae; Drosophila.  
 NX NCBI\_TaxID=7227;  
 [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=2019606; PubMed=10731132;  
 RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,  
 RA Amanatides P.G., Scherer S.E., Li P.W., Hoekins R.A., Galle R.F.,  
 RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,  
 RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,  
 RA Wan K.H., Doyle C., Baxter E.G., Helt G., Champs C., Pfeiffer B.D.,  
 RA Abell J.F., Agbayani A., An H.J., Andrews-Pfannkoch C., Baldwin D.,  
 RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,  
 RA Beeson K.Y., Berens P.V., Berman B.P., Bhandari D., Bolshakov S.,  
 RA Borkova D., Botchan M.R., Bouck J., Brockstein P., Brotler P.,  
 RA Burris K.C., Busam D.A., Butler H., Cadelis E., Center A., Chandra I.,  
 RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,  
 RA DePablo B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,  
 RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,  
 RA Foeller C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,  
 RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,  
 RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,  
 RA Hostin D., Houston K.A., Howland T.J., Wei M.H., Ibegawa C.,  
 RA Jalali M., Kalish F., Kaepfen G.H., Ke Z., Kennison J.A., Ketchum K.A.,  
 RA Kimmel B.E., Kodita C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,  
 RA Laske P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,  
 RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,  
 RA Merkulov G., Mishina N.V., Mobarry C., Morris J., Moshrefi A.,  
 RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,  
 RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pauley J.M.,  
 RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,  
 RA Reinert K., Remington K., Saunders R.D., Scheeler F., Shen H.,  
 RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,  
 RA Spier E., Spredling A.C., Stapleton M., Strong R., Sun E.,  
 RA Svitek R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,  
 RA Wang Z.Y., Wasman D.A., Weinstein G.M., Weissbach J.,  
 RA Williams S.M., Woodgett M., Wu D., Yang S., Yao Q.A., Ye J.,  
 RA Yeh R.F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,  
 RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,  
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;  
 RT "The genome sequence of Drosophila melanogaster.";  
 RL Science 287:2185-2195 (2000).  
 [2]  
 RP SEQUENCE FROM N.A.  
 RA Celniker S.E., Adams M.D., Kronmiller B., Wan K.H., Holt R.A.,  
 RA Evans C.A., Gocayne J.D., Amanatides P.G., Brandon R.C., Rogers Y.,  
 RA Banzon J., An H., Baldwin D., Banzon J., Beeson K.Y., Busam D.A.,  
 RA Carlson J.W., Center A., Champs C., Davenport L.B., Dietz S.M.,  
 RA Dodson K., Dorsett V., Doup L.E., Doyle C., Dresnek D., Fafian D.,  
 RA Ferreira S., Fris E., Galle R.F., Garg N.S., George R.A.,  
 RA Gonzalez M., Houck J., Hoekins R.A., Hostin D., Howland T.J.,  
 RA Ibegawa C., Jalali M., Kruse D., Li P., Mattei B., Moshrefi A.,  
 RA McIntosh T.C., Moy M., Murphy B., Nelson C., Nelson K.A., Nuno J.,  
 RA Pauley J., Parag V., Park S., Patel S., Pfeiffer B.,  
 RA Phouenavong S., Pittman G.S., Puri V., Richards S., Scheeler F.,  
 RA Stapleton M., Strong R., Svitek R., Tector C., Tyler D.,  
 RA Williams S.M., Zaveri J.S., Smith H.O., Venter J.C., Rubin G.M.;  
 RT "Sequencing of Drosophila melanogaster genome.";  
 RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.  
 [3]  
 RP SEQUENCE FROM N.A.  
 RA Miera S., Crosby M.A., Matthews B.B., Bayraktaroglu L., Campbell K.,  
 RA Hradecky P., Huang Y., Kaminker J.S., Prochman S.E., Smith C.D.,

RA Tupy J.L., Bergman C.M., Berman B.P., Carlson J.W., Celniker S.E.,  
 RA Clamp M.E., Drysdale R.A., Emmert D., Frise E., de Grey A.D.N.J.,  
 RA Harris N.L., Kronmiller B., Marshall B., Millburn G.H., Richter J.,  
 RA Russo S., Seale S.M.J., Smith E., Shu S., Smutnack F.,  
 RA Whitfield E.J., Ashburner M., Gelbart W.M., Rubin G.M., Mungall C.J.,  
 RA Lewis S.E.;  
 RT "Annotation of Drosophila melanogaster genome.";  
 RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.  
 RN [4]  
 RP SEQUENCE FROM N.A.  
 RA Flybase;  
 RL Submitted (SEP-2002) to the EMBL/GenBank/DBJ databases.  
 RN [5]  
 RP SEQUENCE FROM N.A.  
 RA Flybase;  
 RL Submitted (FEB-2003) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AB003775; X14261.2;  
 DR GO; GO:0005764; C:lysosome; IEA.  
 DR GO; GO:0005489; F:electron transporter activity; IEA.  
 DR GO; GO:0006118; P:electron transport; IEA.  
 DR GO; GO:0006665; P:sphingolipid metabolism; IEA.  
 DR InterPro; IPR000345; Cyt c heme\_BS.  
 DR InterPro; IPR007856; SapB\_1.  
 DR InterPro; IPR008138; SapB\_2.  
 DR InterPro; IPR008140; SapB\_3.  
 DR InterPro; IPR008373; SapB\_4.  
 DR InterPro; IPR008139; Saposin.  
 DR Pfam; PF05184; SapB\_1\_3.  
 DR Pfam; PF03489; SapB\_2\_4.  
 DR PRINTS; PR01797; SAPOSIN.  
 DR ProDom; PD001732; SapB sub. 2.  
 DR SMART; SM00118; SAPB\_7.  
 DR PROSITE; PS00150; CYTOCHROME\_C\_1.  
 SQ SEQUENCE 876 AA; 97617 MW; 767E16D35ACF52DB CRC64;  
 Query Match 15.7%; Score 432.5; DB 5; Length 876;  
 Best Local Similarity 23.8%; Pred. No. 2.9e-24;  
 Matches 117; Conservative 91; Mismatches 178; Indels 105; Gaps 17;  
 QY 55 LPEDIDKVVYTAGMLKDNATTEBILVLEKTCQWLPKPMMSASCKEIVSYLPVLDI 114  
 DB IPELCEQVYKRLRVLVANTETETFEKQVMEGFCF--QSKGFQDECSIVDQYHVIET 265  
 QY 115 IKGMSRPGEVCSALNLC-----ESLQKHLAEINHQ 145  
 DB IYKSKLDANG-ACCMGICQKNSASMKDVPIMPLVIEPAQVKTTEKLEKH-----EK 319  
 QY 146 KQLESNKIPELDMTEVAPFMANIPLLVYPODGPRSKROP---KDNQDVQCDDCIQWYTDI 202  
 DB KQLGASE-PKRSQGEIID-----MQLPIDHMGANPQALVGEGLCTLCYMLHFI 370  
 QY 203 CFAVNTSTFQALVENVKEE-CDRLGPGNADICRNTISQYSEIATIQMMNH-MQPKICA 260  
 DB QETLTPESTIDE--IKHVENIKCAKLPBGVAGQCRNFVEMVGDVIALVQGLNPRDVC 428  
 QY 261 LVGFCDEYKEMQVLPVAKVASKVIVPALELVIEPIKHEVPAKSDVYCVCEFLVKEVT 320  
 DB LMQKCP--KNLPPK-----EDVEVNPQASDEQD-----PT-----CPLCFAYEQAQ 471  
 QY 321 KLINDNTEKILIAFDKMGSKLPKSLSECCQEVVDYTGSSISLILEVSPVLVSMTH 380  
 DB MKINDNSKDKIKKVLNGLTSHLPNEIKKECVDFVNTYSNELIDMLITDPKPEICQQLK 531  
 QY 381 LCGST--RLPALTVHVTQPPDG-----GF-----CEV 405  
 DB LCPRTTALMDRLRSLDEVDVDEGDKSSSEIISNDIESLEELPQALFDPGFPAAENCI 591  
 QY 406 CKKVLGVLDNRNLEKNSKQKQKILALEKQCSFLPDYPQKODQFVAYEYVLLIILVWMD 465  
 DB CEELVKTLEKRMGKHPRDSIKHILLESQDRMKKPNNTCHVYIDYXGXKIDLLEKMD 651  
 QY 466 PSFVCLTKIGAC 476



Db 652 PKLICTELGMC 662

## RESULT 13

AC Q95X02 PRELIMINARY; PRT; 458 AA.  
 ID Q95X02  
 RT "Pore-forming peptides of Naegleria fowleri."  
 DE Naegleria fowleri.  
 GN NP-B.  
 OS Naegleria fowleri.  
 OC Eukaryota; Heterolobosea; Schizopyrenida; Vahlkampfiidae; Naegleria.  
 OX NCBI\_Taxid=5763;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Nickel R., Leipzig M.;  
 RT Submitted (OCT-1999) to the EMBL/Genbank/DBJ databases.  
 DR EMBL; AF196309; AAL0158.1; -  
 DR GO; GO:0005764; C:lysosome; IEA.  
 DR GO; GO:0006665; P:sphingolipid metabolism; IEA.  
 DR InterPro; IPR007856; SApB\_1.  
 DR InterPro; IPR008138; SApB\_2.  
 DR InterPro; IPR008140; SApB\_sub.  
 DR InterPro; IPR008373; SaposinB.  
 DR Pfam; PF05184; SApB\_1; 5.  
 DR Pfam; PF03489; SApB\_2; 5.  
 DR PRINTS; PR01797; SAPOSIN.  
 DR ProDom; PD001732; SApB\_sub; 4.  
 DR SMART; SM00118; SApB; 5.  
 FT NON\_TER  
 SQ SEQUENCE 458 AA; 50452 MW; 433F60CDE64C904 CRC64;

Query Match 12.5%; Score 343.5; DB 5; Length 458;  
 Best Local Similarity 25.6%; Pred. No. 7.4e-18;  
 Matches 115; Conservative 69; Mismatches 186; Indels 79; Gaps 16;

QY 36 CGAVHGC---LQTVNKKPTVKSIPCDICQDVTAAGMLKDNATEELVLYLEKTCMDLP 92  
 Db 69 CSQVVICNGSAMAAVAPRAENSGICNMCOQLVTVQENNVESNDTITMLEKLEQCVSVIP 128  
 QY 93 KPNMSASKEIYDSYSLPVLIDIKGMSRPG-EVCSALNLCESLQKHLAELNHQQLBSN 151  
 Db 129 -GQYSALCTYAVBOYLPIFIHQV--EKOPPALITICQDVHLCSAQ----- 170  
 QY 152 KIPELDMTEVVAFFMANIPLLLYPODGRSKPQPKDNGVQDCIQMTDIOQAVRTNST 211  
 Db 171 -----AAPVV-----QQQAAELCPICKANGFLTKI--NNV 201  
 QY 212 FVQALVEHVKEECRLGPGMADICKNYISQYSEIAIQMMHMQPEIKCALVGFCEVEM 271  
 Db 202 DVNAVKQQLERACSF--QVPDCQIVDKAQLA-QDLQTEDAQITCSTV-----VVC 252  
 QY 272 PMQITVPAKVASKNVIALPELVEPIKGEVPAKSDVYCEVGFYKVEYTKLIDNNKTEKE 331  
 Db 253 PKQGVV-----TENPKKF--LEAKDSKYCPTCLOITKYLEDLIVSDITVNE 297  
 QY 332 ILDAFDKMSCKLPSLSSEECQEVVDYTGSSILSLLEEVSPELVCSMLHCSGTRLPALT 391  
 Db 298 IIKLADAGCARL-GALBSLCKKFFVPLAVDELKLLLEKLPQKVCSTLMKCDAAELIKLA 356  
 QY 392 VHVTPKDGGFCEVCKKLVGLDRLNLEKNSYKQELIALLEKGC-SFLPDPYQKQDDQFVA 450  
 Db 357 L-APQAAQGTWCLACEYVISVADNMLINNTQOSVGNFLDKVCOEFVPSITQSOCTALVN 415  
 QY 451 EYEPVLIETL-VEVWDPSFVCLKIGACPS 478  
 Db 416 QYEAQLIQLFESKVFNPQTVCKAIGVCS 444

## RESULT 14

AC Q9BKML PRELIMINARY; PRT; 484 AA.  
 ID Q9BKML  
 RT "Pore-forming peptides of Naegleria fowleri."  
 DE Naegleria fowleri.  
 GN PRONP-B.  
 OS Naegleria fowleri.  
 OC Eukaryota; Heterolobosea; Schizopyrenida; Vahlkampfiidae; Naegleria.  
 OX NCBI\_Taxid=5763;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Nickel R., Benkert C., Jacobs T., Marti T., Marciano-Cabral F.,  
 RA Leipzig M.;  
 RT Submitted (MAY-1999) to the EMBL/Genbank/DBJ databases.  
 DR EMBL; AF154047; AAK21659.1; -  
 DR GO; GO:0005764; C:lysosome; IEA.  
 DR GO; GO:0006665; P:sphingolipid metabolism; IEA.  
 DR InterPro; IPR007856; SApB\_1.  
 DR InterPro; IPR008138; SApB\_2.  
 DR InterPro; IPR008140; SApB\_sub.  
 DR InterPro; IPR008373; SaposinB.  
 DR Pfam; PF05184; SApB\_1; 5.  
 DR Pfam; PF03489; SApB\_2; 5.  
 DR PRINTS; PR01797; SAPOSIN.  
 DR ProDom; PD001732; SApB\_sub; 4.  
 DR SMART; SM00118; SApB; 5.  
 SQ SEQUENCE 484 AA; 53237 MW; 63DA3AAE7CA578B0 CRC64;

Query Match 12.5%; Score 343.5; DB 5; Length 484;  
 Best Local Similarity 25.6%; Pred. No. 8e-18;  
 Matches 115; Conservative 69; Mismatches 186; Indels 79; Gaps 16;

QY 36 CGAVHGC---LQTVNKKPTVKSIPCDICQDVTAAGMLKDNATEELVLYLEKTCMDLP 92  
 Db 95 CSQVVICNGSAMAAVAPRAENSGICNMCOQLVTVQENNVESNDTITMLEKLEQCVSVIP 154  
 QY 93 KPNMSASKEIYDSYSLPVLIDIKGMSRPG-EVCSALNLCESLQKHLAELNHQQLBSN 151  
 Db 155 -GQYSALCTYAVBOYLPIFIHQV--EKOPPALITICQDVHLCSAQ----- 196  
 QY 152 KIPELDMTEVVAFFMANIPLLLYPODGRSKPQPKDNGVQDCIQMTDIOQAVRTNST 211  
 Db 197 -----AAPVV-----QQQAAELCPICKANGFLTKI--NNV 227  
 QY 212 FVQALVEHVKEECRLGPGMADICKNYISQYSEIAIQMMHMQPEIKCALVGFCEVEM 271  
 Db 228 DVNAVKQQLERACSF--QVPDCQIVDKAQLA-QDLQTEDAQITCSTV-----VVC 278  
 QY 272 PMQITVPAKVASKNVIALPELVEPIKGEVPAKSDVYCEVGFYKVEYTKLIDNNKTEKE 331  
 Db 279 PKQGVV-----TENPKKF--LEAKDSKYCPTCLOITKYLEDLIVSDITVNE 323  
 QY 332 ILDAFDKMSCKLPSLSSEECQEVVDYTGSSILSLLEEVSPELVCSMLHCSGTRLPALT 391  
 Db 324 IIKLADAGCARL-GALBSLCKKFFVPLAVDELKLLLEKLPQKVCSTLMKCDAAELIKLA 382  
 QY 392 VHVTPKDGGFCEVCKKLVGLDRLNLEKNSYKQELIALLEKGC-SFLPDPYQKQDDQFVA 450  
 Db 383 L-APQAAQGTWCLACEYVISVADNMLINNTQOSVGNFLDKVCOEFVPSITQSOCTALVN 441  
 QY 451 EYEPVLIETL-VEVWDPSFVCLKIGACPS 478  
 Db 442 QYEAQLIQLFESKVFNPQTVCKAIGVCS 470

RESULT 15  
 ID Q35489 PRELIMINARY; PRT; 378 AA.

```

AC 035489;
DT 01-JAN-1998 (TrEMBLrel. 05, Created)
DT 01-JAN-1998 (TrEMBLrel. 05, last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, last annotation update)
DE Surfactant protein-B.
OS Cavia porcellus (Guinea pig).
OC Eukaryota; Metazoa; Chordata; Cranialata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Hystricognathi; Caviidae; Cavia.
OX NCBI_TaxID=10141;
RN [1]
RP SEQUENCE FROM N.A.
RA Bingle C.D., Yuan H.T., Gowan S.;
RT "Guinea pig surfactant protein-B is differentially polyadenylated.";
RL Submitted (NOV-1997) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF033190; AAB87086.1; -.
DR HSSP; P07988; 1DFW.
DR GO; GO:0005764; C:lysosome; IEA.
DR GO; GO:0007585; P:respiratory gaseous exchange; IEA.
DR GO; GO:0006665; P:sphingolipid metabolism; IEA.
DR InterPro; IPR003119; Sapa.
DR InterPro; IPR007856; Sapa.1.
DR InterPro; IPR008138; Sapa.2.
DR InterPro; IPR008140; Sapa_sub.
DR InterPro; IPR008373; Saposin.
DR InterPro; IPR008137; SaposinB.
DR Pfam; PF02199; SAPA; 1.
DR Pfam; PF05184; Sapa_1; 2.
DR Pfam; PF03489; Sapa_2; 3.
DR PRINTS; PR01797; Saposin.
DR ProDom; PD001732; Sapa_sub; 1.
DR ProDom; PD008002; Surfactant_B; 1.
DR SMART; SM00162; SAPA; 1.
DR SMART; SM00118; Sapa; 3.
SQ
SEQUENCE 378 AA; 41677 MW; E3DA2E237ED401DA CRC64;

Query Match      12.4%; Score 342.5; DB 11; Length 378;
Best Local Similarity 22.6%; Pred. No. 76-18;
Matches 111; Conservative 67; Mismatches 174; Indels 139; Gaps 15;

QY 1 LASLIGALLAGPVILGKECTRGSAVWQNVKTASDCGAVKHQLOTVWNNKPTVKSJPCDIC 60
DB 13 LPMLCGPGAAQTTSILACSGQPFKFCQSLBQALGCRALGHQLOEWHGVGADDL-CQEC 71
QY 61 KDVVTAAGDMKDNATEBEELVLEKTCDMLEPKRMASCKEIVDSYLPVLDITIKGEMS 120
DB 72 EDIVAILTKMAKEVIFQKTKRFLERECDELPLKILVPRCHSVLETYFPLVIDYFQSHIT 131
QY 121 RPEGVCSALNLCESLOKHLAEINHQKLESNKIPELDMTEVAVPFMANIPLLLYPDGPR 160
DB 132 -PKTICWNLGLCQPRQ-----PD----- 148
QY 181 SKPPQKNGDVQCQDCIQWNTDIOIAVRTNSTFVQALVEHVKEECRLPGWADICKNYIS 240
DB 149 --PQPE-----PGMPSPLPNSLS 164
QY 241 QYSEIATQMNMHMQPKICALVGFCDVEKEMPQTLVPARY-ASKNVITPALVELVPIKKA 299
DB 165 ETSF-----DLVPPRLPRALSVWPGPH--TODLSEQ 194
QY 300 EVPAKSDVYCEVCEFLVKEVTKLIDNNKTEKELD-AFDKMSCKLPKSLSECEVVDVY 358
DB 195 QFPIPLP-YGRICKTKLKRQAMT-----PKGVLAAMAVAOCHVPLVAGGICQCLAERY 248
QY 359 GSSLSILIEVSPELVCSMLHLS-----GTRLPALTVAHTQPKGGFCEVCKLVGYL 413
DB 249 TVLLDLALSHLPLQVLGVLRCSDMSAGLVLPALSLSEGWLPQSECDLCLSVT--- 305
QY 414 DRNLKSTKQETLALAEKSGSFLPDYQKQ-CDQFAVEYEPVLIIEVWMPSPVCLK 472
DB 306 --TRAWNSSELARQAMHQAC--LSSPLDRQCKQVEEQYTPQLALDEPRGDPRTTQCA 361
QY 473 IGACPSAHKPL 483

```

Db :|||  
362 LGVCRGANPL 372

Search completed: May 5, 2004, 13:33:41  
Job time : 52.6524 secs